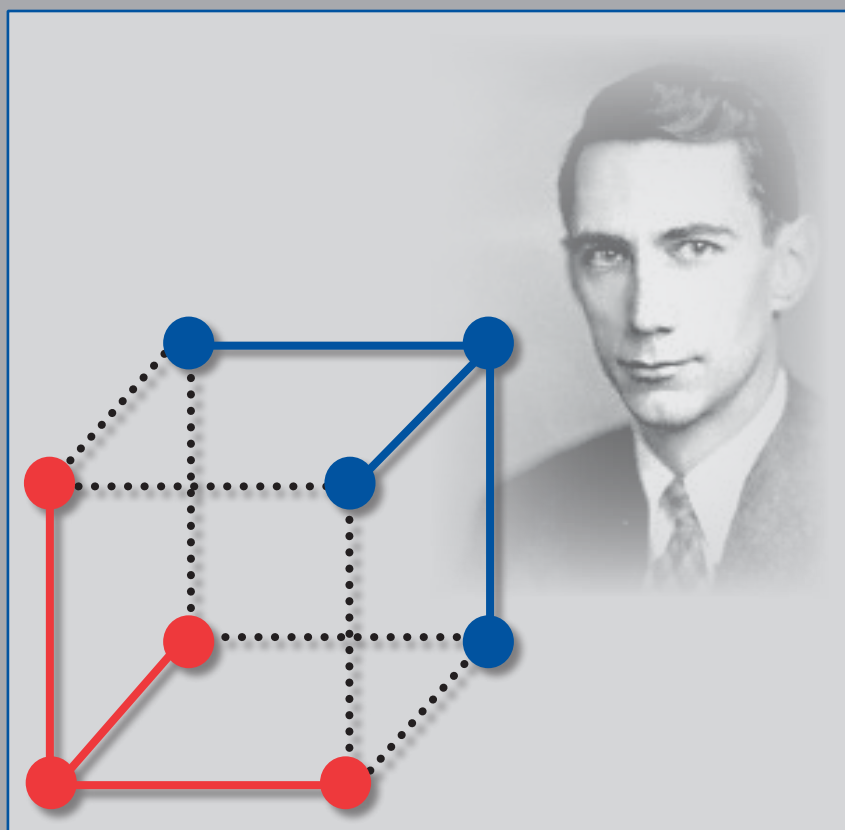
Marzio Barbero - Natasha Shpuza

Errori e Parità

1

Serie di articoli, pubblicati in più numeri di Elettronica e Telecomunicazioni, trattano e approfondiscono una singola tematica. Lo scopo dell'iniziativa **LeMiniSerie** è di raccogliere tali articoli, con una veste tipografica unitaria che ne faciliti la consultazione e apportando correzioni e aggiornamenti ritenuti opportuni.

Questo è il primo di questi volumi.

Errori e Parità raccoglie tre articoli pubblicati a nei numeri di aprile e agosto 2008 e agosto 2010.

LeMiniSerie sono una iniziativa del
Centro Ricerche e Innovazione Tecnologica della
www.crit.rai.it

In copertina:
ritratto di *Claude E. Shannon* e illustrazione relativa alla *distanza di Hamming*.

Nel 1948 Claude E. Shannon pubblica un articolo destinato a rivoluzionare la tecnologia alla base delle telecomunicazioni digitali. In tale articolo enuncia quello che è noto come Limite di Shannon, fornendo una relazione fra la capacità massima, in termini di bir-rate, e la larghezza di banda di un canale rumoroso. La capacità può essere raggiunta grazie all'uso di codici per la correzione degli errori e Shannon indica nel codice di Hamming un esempio di codice efficiente.

L'uso dei bit di parità nel codice di Hamming è adottato come esempio, nel primo capitolo, per facilitare la comprensione delle tecniche che consentono di rivelare e correggere gli errori introdotti dal canale, di trasmissione o registrazione. Segue un approfondimento sul ruolo di tali tecniche, ed in particolare con l'uso dei codici di Reed-Solomon, nei sistemi utilizzati per la diffusione e registrazione del segnale televisivo, professionali e consumer, quali il CD e il DVD.

Il secondo capitolo, partendo dal contributo fondamentale di Shannon, ricorda brevemente le tappe principali dello sviluppo di teorie e tecniche per la correzione degli errori e illustra il ruolo fondamentale di queste tecniche per lo sviluppo delle esplorazioni spaziali. Un ruolo determinante è quello di Andrew Viterbi e di Irving Reed.

Il terzo capitolo illustra i sistemi basati sulla concatenazione di codici convoluzionali e codici Reed Solomon e la decodifica di Viterbi e il ruolo di tale schema non solo in campo spaziale, ma anche nello sviluppo della prima generazione dei sistemi di diffusione e distribuzione digitale televisiva, via satellite e terrestre. L'avvento dei turbo codici, nel 1993, e la riscoperta dei codici LDPC, ovvero dell'impiego dei bit di parità in codici a blocco molto lunghi, già proposta da Robert Gallager nel lontano 1962, comporta un salto di qualità inaspettato e, praticamente, il raggiungimento del limite di Shannon. Tutti i più recenti sistemi di telecomunicazione, a partire dalla seconda generazione degli standard per la diffusione e distribuzione dei segnali televisivi e dai nuovi sistemi di protezione dei dati per la registrazione, adottano schemi basati su LDPC.

Torino, ottobre 2010

M. Barbero e N. Shpuza Errori e Parità

Acronimi e sigle

3GPP	3rd Generation Partnership Project
APP	A Posteriori Probability
ASI	Agenzia Spaziale Italiana (www.asi.it)
ATM	Asyncronous Transfer Mode
ATSC	Advanced Television Systems Committee (www.atsc.org)
AWGN	Additive White Gaussian Noise
BCH	Bose, Chaudhuri, Hocquenghem (codice)
BER	Bit Error Rate
BIS	Burst Indication Subcode
BVD	Big Viterbi Decoder
CCSDS	Consultative Committee for Space Data Systems (www.ccsds.org)
CD	Compact Disc
CRC	Cyclic Redundancy Check
DSP	Digital Signal Processor
DTMB	Digital Terrestrial Multimedia Broadcast
DVB **-S** **-RCS** **-RCT** **-S2** **-T2**	Digital Video Broadcasting, (www.dvb.org) - Satellite - Return Channel over Satellite - Return Channel over Terrestrial - Satellite (new generation) - Terrestrial (new generation)
DVD	Digital Versatile Disc (www.dvdforum.org)
ECC	Error Correcting Code
ESA	European Space Agecy (www.esa.int)
FEC	Forward Error Correction
FLUTE	File Delivery over Unidirectional Transport
G.hn	home network
HDTV	High Definition TeleVision
IRA	Irregular Repeat Accumulate
JPL	Jet Propulsion Laboratory (www.jpl.nasa.gov)
LDC	Long Distance Code
LDPC	Low-Density Parity-Check
LSI	Large Scale Integration
LT	Luby Transform (codice)
MDS	Maximum Distance Separable
MBMS	Multimedia Broadcast/Multicast Services
MDS	Maximum Distance Separable
MPEG	Motion Picture Expert Group
MRO	Mars Reconnaisance Orbiter
NASA	National Aereonautics and Space Administration (www.nasa.gov)
PDM	Professional Disc Media
PLL	Phase Lock Loop
RAID	Redundand Array of Inexpensive Disks oppure Redundand Array of Independent Disks
RAPTOR	RAPid TORnado
RM	Reed Muller (codice)
RS	Reed Solomon (codice)
SHV	Super Hi-Vision
SISO	Soft-In-Soft-Out
SoC	System on Chip
SOVA	Soft-Output Viterbi Algorithm
UMTS	Universal Mobile Telecommunications System
VA	Viterbi Algorithm
VLSI	Very Large Scale Integration

1 I codici: da Hamming a Reed-Solomon

Testo e figure tratte da "Rivelazione, Correzione e Mascheramento degli Errori - Parte I" di Marzio Barbero, Natasha Shpuza , Elettronica e Telecomunicazioni, Aprile 2008

1. Correggere gli errori per garantire la qualità

Il successo dei sistemi di comunicazione e memorizzazione digitali dei segnali audio e video è essenzialmente dovuto all'impiego in tali campi delle stesse tecnologie che sono state sviluppate per i calcolatori elettronici, e che, grazie alla loro evoluzione, consentono oggi di elaborare con velocità adeguate gli elevati flussi di dati associati alle informazioni audiovisive.

Le informazioni audio e video codificate secondo gli standard digitali [1,2] sono di tipo binario e ciò ha reso economicamente molto vantaggioso lo sviluppo degli apparati digitali rispetto a quelli che operavano in ambito analogico. Infatti i segnali binari presentano solo due stati, o valori, in termini di corrente, tensione o magnetizzazione.

La trasformazione dei segnali video e audio in dati numerici consente inoltre di utilizzare tecniche atte a contrastare il degradamento che subisce l'informazione quando i segnali transitano sul canale hertziano o su cavo oppure vengono memorizzati. I canali, infatti, possono introdurre una serie di disturbi (rumore, interferenze, distorsioni, echi) che, combinandosi con il segnale utile, causano una non corretta identificazione dell'informazione in ricezione o in lettura.

Il flusso dati in ricezione non coincide quindi con quello trasmesso, ma a causa del canale, è caratterizzato da un certo tasso di errore (BER). Nel caso in cui i segnali audio e video vengano codificati utilizzando tecniche per la riduzione della ridondanza [3,4] è essenziale che praticamente tutti i dati trasmessi o memorizzati pervengano al decodificatore inalterati: solo in tal caso l'informazione originaria può essere riprodotta fedelmente, e senza degradamenti percettibili della qualità.

Tale condizione di "quasi assenza di errori" è quantizzabile, in termine di BER, con valori estremamente bassi, dell'ordine 10^{-10} (corrispondente ad esempio, nel caso di un flusso di 5 Mbit/s, a circa 2 bit errati ogni ora): solo con tassi di errore di quest'ordine di grandezza le immagini o l'audio codificate con le moderne tecniche di compressione possono essere riprodotte senza che i difetti introdotti dal canale siano considerati fastidiosi dall'utente, grazie alla bassissima frequenza con cui si presentano.

Una probabilità di errore così bassa può essere fornita, fra i canali normalmente utilizzati, solo dalla fibra ottica. La maggior parte degli altri canali è caratterizzata da BER superiori di diversi ordini di grandezza (arrivando anche a valori di $10^{-3} \sim 10^{-4}$): i sistemi digitali sarebbero poco competitivi, se non fossero stati sviluppati sistemi atti a rilevare e correggere gli errori.

E' quindi stata fondamentale l'evoluzione, nel corso degli ultimi due decenni, delle tecniche di correzione basata sui codici; tecniche che hanno potuto trovare una applicazione pratica grazie all'incremento in termini di velocità e di capacità di memorizzazione di cui dispongono le attuali unità

di elaborazione (microprocessori e DSP), in grado di effettuare in tempo reale le numerose e, a volte, complesse operazioni necessarie per la decodifica.

E' quindi la progettazione di sistemi che integrano in modo ottimale le tecniche di compressione delle informazioni audio e video a quelle di protezione dagli errori e di modulazione, alla base dello spettacolare progresso tecnico della diffusione di immagini e suoni: aumento della densità di memorizzazione su supporti magnetici e ottici, aumento dell'efficienza spettrale delle comunicazioni, aumento dell'area di copertura dei servizi radiotelevisivi, diminuizione della potenza in trasmissione, via satellite e terrestre.

Si intende qui offrire una panoramica, a partire dalle prime ad essere introdotte in campo televisivo alla fine degli anni '70, delle tecniche per la rivelazione e correzione degli errori utilizzate nei sistemi per la diffusione e memorizzazione dei segnali video e audio digitali, pur evitando di approfondire le teorie alle basi dei codici, teorie che spesso implicano l'impiego di strumenti matematici complessi.

2. La rivelazione degli errori: il bit di parità

Il codice più semplice adottato per la rivelazione degli errori (*error detection*) introdotti da un canale di comunicazione è costituito dall'aggiunta di un solo bit di ridondanza per ciascuna delle parole in cui sono organizzati i dati trasmessi.

Il bit di parità, pari o dispari, consente di rilevare la presenza di errori, ma non consente di correggerli (figura 1).

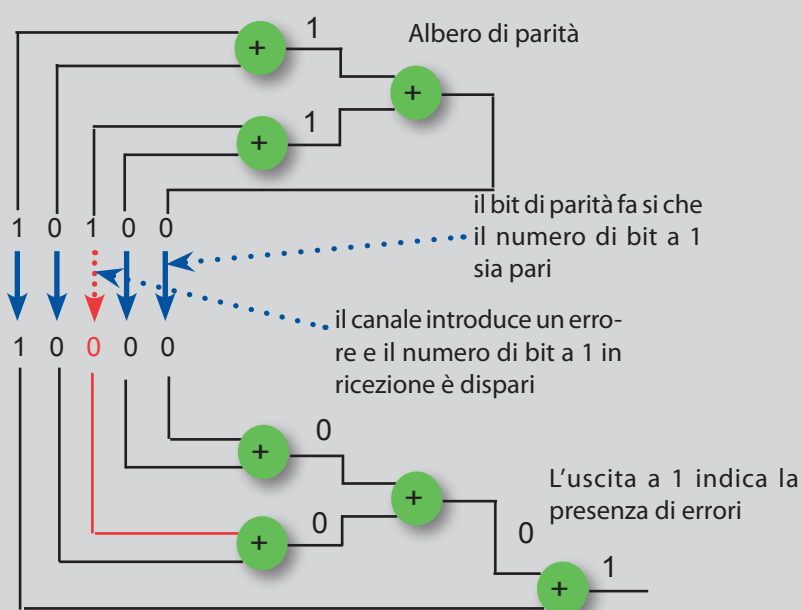
E' utile in presenza di probabilità di errore molto bassa, per verificare che il canale non presenti problemi di trasmissione, o nel caso in cui, avendo a disposizione un canale bidirezionale, sia possibile richiedere la ritrasmissione della parola errata. E' efficace anche nel caso in cui l'informazione venga ritrasmessa ciclicamente, come avviene nel caso del servizio Televideo (figura 2) .

Tabella XOR

A	B	C
0	0	0
0	1	1
1	0	1
1	1	0

Fig. 1 - Vi sono due tipi di parità: il bit di parità pari e il bit di parità dispari.

Nel caso di bit parità pari, come in questo esempio basato su una parola di 4 bit di informazione, si aggiunge un bit 1 nel caso in cui la parola abbia un numero dispari di bit 1 oppure viene aggiunto uno 0 se tale numero è pari. Ciò si ottiene mediante semplici operazioni di somma binaria (XOR). Nell'esempio si trasmettono 5 bit. In ricezione la presenza, ma non la posizione, di un singolo bit errato è rivelata con modalità altrettanto semplici.

Nel caso di bit di parità dispari si fa sì che il numero di bit 1 trasmessi sia sempre dispari.

Una riga di cancellazione di quadro del segnale televisivo analogico vista all'oscilloscopio: si vedono: il sincronismo di riga, il burst colore e i dati teletext.

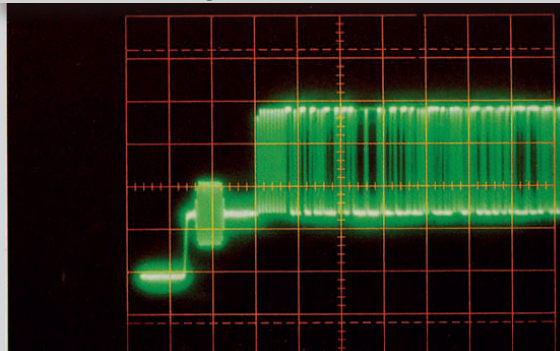



Fig. 2 - Nel sistema teletext, utilizzato per il servizio Televideo in Italia, i dati (pagine di 24 righe di 40 caratteri ciascuna) sono trasmessi sulle righe di cancellazione di quadro del segnale video della TV analogica.

I caratteri di testo o di controllo sono codificati come parole da 8 bit (byte o ottetto) di cui 7 bit sono di informazione e uno è il bit di parità.

Grazie al bit di parità è possibile individuare la presenza di un errore nel carattere ricevuto, in tal caso nella memoria di pagina viene memorizzato un indicatore (*flag*) e in visualizzazione lasciata vuota la posizione corrispondente.

In questo modo, nella successiva ricezione della stessa pagina (la trasmissione delle pagine televideo è ciclica), a meno che venga nuovamente rivelata la presenza di errore, il carattere può essere memorizzato e visualizzato. Anche nei casi in cui il canale sia affetto da una elevata probabilità di errore, il bit di parità risulta efficace come sistema per la rivelazione e correzione degli errori, grazie alla ridondanza insita nella ritrasmissione ciclica dell'informazione.

(foto tratte da [5]).

Immagine di prova Televideo (Monoscopio Televideo) contiene tutti i caratteri di controllo e quelli visualizzabili.

In presenza di errori, essi sono facilmente identificabili nella pagina di test, che viene riacquisita ad ogni ciclo.

3. Codici a blocco e distanza di Hamming

In un codice a blocco i simboli (normalmente simboli binari) da codificare sono organizzati in parole di lunghezza finita prima di essere avviati al canale. Lo scopo della codifica è quello di consentire la localizzazione e la correzione di uno o più simboli, fra quelli costituenti la parola, che eventualmente pervengano errati al decodificatore.

Alla singola parola costituita da *k* simboli si aggiunge un numero predefinito di simboli in modo che la lunghezza totale della parola trasmessa, o memorizzata, diventi di *n* simboli. L'algoritmo di codifica aggiunge dunque *n-k* simboli di controllo, che costituiscono la ridondanza necessaria a permettere la localizzazione e correzione degli errori introdotti sul canale.

Nel caso in cui i simboli siano binari, cioè bit che possono assumere solo i valori 0 e 1, le combinazioni che possono essere generate dal codificatore sono 2^k, mentre le combinazioni possibili, e che possono pervenire al decodificatore, sono tutte le 2^n combinazioni degli *n* bit.

Un codice a blocco molto semplice è quello riprodotto in figura 3: in questo caso la decodifica è possibile grazie alla ripetizione del bit trasmesso e all'adozione, in fase di decodifica, di un criterio maggioritario, di massima verosimiglianza.

La stessa figura illustra il concetto di distanza (distanza di Hamming[Nota 1]) fra due parole: è il numero di posizioni per cui i simboli corrispondenti differiscono fra loro.

La distanza fra due parole rappresenta il numero di "errori" che trasformano una parola nell'altra. Se si calcolano i valori delle distanze fra tutte le coppie di parole codice che la sorgente può emettere e si individua il valore minimo, questo è detto distanza minima d_{min} e si può dimostrare che il numero *e* di errori correggibili in una parola è pari a $e = (d_{min}-1)/2$.

Un codice a blocco è quindi caratterizzato dai tre valori (n, k, d_{min}), la ridondanza è legata al numero di simboli aggiunti per la protezione, cioè *n-k*.

Non è possibile avere un codice che comprenda un elevato numero di parole di codice (valore di *k* alto), che corregga molti errori (d_{min} alto) e le cui parole di codice siano corte (*n* basso): occorre individuare il miglior compromesso fra i suddetti tre parametri.

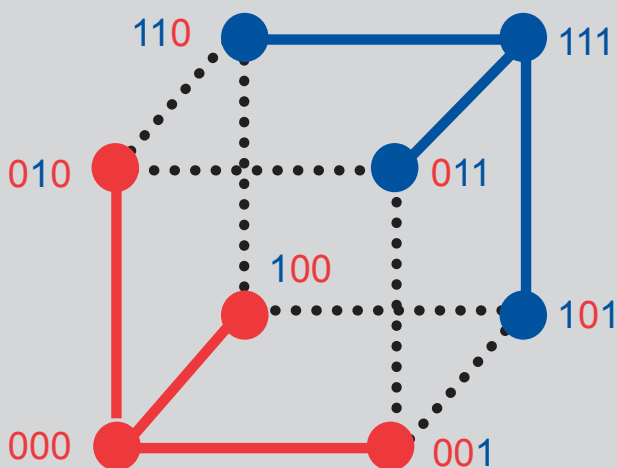
Le tecniche di correzione degli errori che consentono l'identificazione diretta degli errori, senza richiedere la trasmissione ciclica o la ritrasmissione dell'informazione, sono definiti FEC.

Nota 1 - Richard Hamming si veda pag. 21.

Fig. 3 - In un codice con ripetizione il bit trasmesso è ripetuto *n* volte.

Ad esempio con un fattore di ripezione 3, le combinazioni trasmesse sono 000 e 111, in corrispondenza ai due simboli *0* e *1*, e si ottiene un codice a blocco (*n*=3, *k*=1).

In ricezione tutte le combinazioni di tre bit sono possibili, ma, assumendo che il canale non introduca più di un errore in una parola, si sceglie un criterio di massima verosimiglianza e quindi si decide che il simbolo trasmesso è quello meno distante da quello ricevuto.

La distanza *d* (detta distanza di Hamming) fra due parole è il numero di posizioni per cui i simboli corrispondenti differiscono fra loro.

In questo cubo ad ogni vertice corrisponde una delle 8 possibili parole da 3 bit. I vertici rossi hanno distanza ≤1 dalla parola 000, analogamente i vertici blu hanno distanza ≤1 dalla parola 111.

4. Codice di Hamming

Un codice a blocco semplice è quello che Hamming ha proposto nel 1950 [6] ed è utilizzato, ad esempio, nel teletext per identificare il numero di magazzino (in pratica la cifra delle centinaia che identifica la pagina Televideo), il numero di riga della pagina e le due cifre relative a decine e unità del numero di pagina.

Tali informazioni sono organizzate in gruppi di 4 bit e tali bit sono ritenuti molto importanti, perché associati alla corretta identificazione di pagina e riga. Una loro erronea decodifica implica la visualizzazione di una pagina non voluta o l'incorretta disposizione delle righe all'interno della pagina: in entrambi i casi il malfunzionamento è evidente all'utente, almeno per tutto il periodo intercorrente prima della ritrasmissionie ciclica della pagina selezionata. Le informazioni associate ai caratteri visualizzabili sono invece ritenute meno importanti e, come si è visto, è ritenuto sufficiente l'uso del bit di parità per la rivelazione degli errori, la loro correzione è affidata alla ridondanza insita nella ritrasmissione ciclica delle pagine.

Questi gruppi di 4 bit sono protetti mediante un codice di Hamming (7,4,d_{min}=3), in grado di correggere errori singoli grazie a 3 bit di ridondanza. I codici di Hamming sono codici lineari e possono essere calcolati mediante l'uso di matrici (figura 4).

Si osservi che i 3 bit di ridondanza consentono di realizzare 2^3=8 configurazioni (denominate sindromi) di cui 7 sono utilizzate in ricezione per individuare quale dei 7 bit è affetto da errore e quindi correggerlo.

Il valore *n-k* è strettamente legato al numero di sindromi, e di conseguenza al numero di posizioni dell'eventuale errore identificabile: ad esempio il codice di Hamming (15,11) può utilizzare i 4 bit di ridondanza per individuare 16 sindromi, di cui quelle non nulle sono sufficienti a identificare la

$$\begin{vmatrix} 1\,0\,0\,1 \end{vmatrix} \begin{vmatrix} 1\,0\,0\,0\,0\,1\,1 \\ 0\,1\,0\,0\,1\,0\,1 \\ 0\,0\,1\,0\,1\,1\,0 \\ 0\,0\,0\,1\,1\,1\,1 \end{vmatrix} = \begin{vmatrix} 1\,0\,0\,1\,1\,0\,0 \end{vmatrix}$$

in trasmissione si calcola: x = u G
il canale può aggiungere errori: y = x + e
in ricezione si calcola la sindrome: s = y H^T

$$\begin{vmatrix} 0\,0\,0\,1\,0\,0\,0 \end{vmatrix}$$

$$\begin{vmatrix} 1\,0\,0\,0\,1\,0\,0 \end{vmatrix} \begin{vmatrix} 0\,1\,1 \\ 1\,0\,1 \\ 1\,1\,0 \\ 1\,1\,1 \\ 1\,0\,0 \\ 0\,1\,0 \\ 0\,0\,1 \end{vmatrix} = \begin{vmatrix} 1\,1\,1 \end{vmatrix}$$

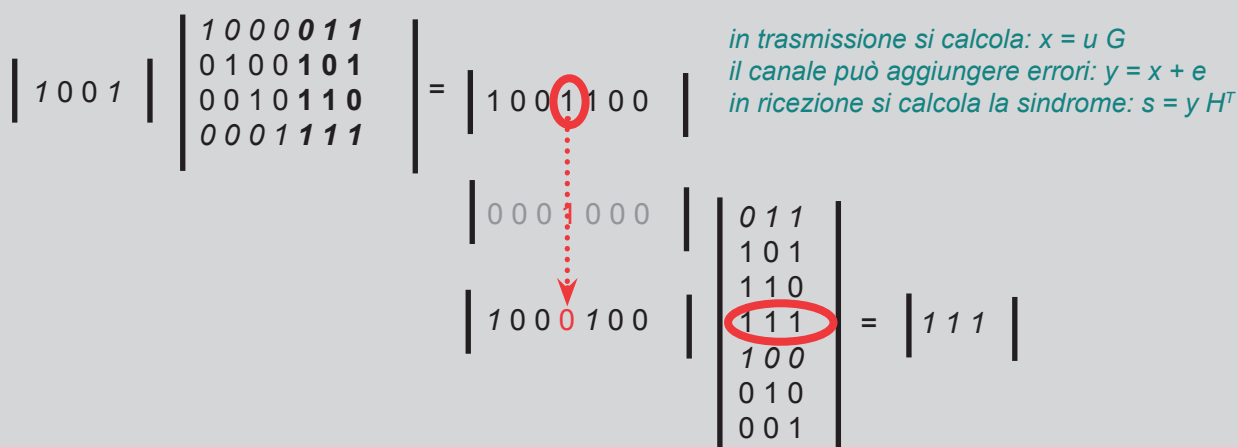



Fig. 4 - Esempio di funzionamento del codice di Hamming (7,4).
A partire dai 4 bit generati dalla sorgente si calcolano 3 bit di ridondanza moltiplicando il vettore *u* di 4 bit (nell'esempio 1001) per la matrice *G* che consente di generare la parola di codice. Il calcolo consiste nel sommare (XOR) le righe di *G* corrispondenti ai bit 1 della parola (vettore) *u:* in questo esempio la prima e la quarta riga, in carattere corsivo in figura per facilitarne la comprensione.
La parola trasmessa risuta cosi essere il vettore *x* di sette bit.
Nell'esempio si suppone che il canale crei un errore sul quarto bit, che da 1 diventa quindi 0. La parola ricevuta è *y*, che differisce da quella trasmessa *x* per il bit in quarta posizione.
In ricezione si calcola un vettore detto sindrome (*s*), sommando le righe della matrice H^T (*H* è la matrice di controllo delle parità i cui elementi sono ricavati a partire dalla matrice generatrice *G*) corrispondenti agli 1 di *y* (in figura gli elementi delle righe di H^T che, sommati, danno origine alla sidrome, sono in carattere corsivo).
Se *s* è formato da tutti bit 0, si conclude che la parola ricevuta non è affetta da errore; negli altri casi, il valore di *s* indica quale è la posizione del bit errato.
In questo esempio *s*=[1 1 1] che corrisponde alla quarta riga di H^T, ovvero al quarto bit, che quindi viene riportato ad 1, correggendo l'errore introdotto dal canale.

posizione dell'errore eventualmente presente in uno degli n=15 bit.

I codici correggono gli errori, in questo caso errori singoli, indipendentemente dal fatto che ne siano affetti i bit di informazione o quelli di ridondanza.

5. Effetti del superamento della capacità di correzione del codice

Nel caso in cui il numero di errori superi la capacità correttiva del codice, in ricezione, anziché recuperare l'informazione corretta, vengono introdotti ulteriori errori (figura 5). Se la capacità di correzione è insufficiente si ha propagazione degli errori, in ricezione si passa rapidamente da una condizione ottimale (tutti gli errori sono recuperati e l'informazione risulta integra) ad una condizione di non funzionamento.

Questo fenomeno spiega le caratteristiche dei sistemi digitali confrontati con quelli analogici: i sistemi digitali sono caratterizzati da un rapido degradamento delle condizioni di servizio qualora si abbia un superamento della capacità di correzione degli errori, a differenza dei sistemi analogici che presentano un degradamento graduale delle prestazioni.

Occorre quindi progettare il sistema di correzione tenendo presenti le caratteristiche del canale e di criticità delle informazioni trasmesse, scegliendo il miglior compromesso fra ridondanza aggiunta e capacità di correzione o rilevazione degli errori.

Tornando all'esempio del sistema teletext, è ritenuto importante non solo proteggere gli indirizzi di pagina e riga mediante un codice, ma si è anche voluta limitare la probabilità di visualizzare una pagina diversa da quella selezionata a causa della presenza di errori in numero superiori all'errore singolo. Pertanto, ai 7 bit del codice $(7,4,d_{min}=3)$ è stato aggiunto un ulteriore bit di parità e la distanza minima passa a $d_{min}=4$: sono così rivelati gli errori di ordine pari, in particolare gli errori doppi.

In ogni caso gli errori di ordine dispari superiore a 1 (3, 5 e 7 errori) non possono essere rivelati e quindi in presenza di un così elevato numero di errori la capacità di correzione e rivelazione del codice viene superata.

I sistemi di correzione degli errori si basano sull'assunto che il BER sia basso e che gli errori siano distribuiti in modo uniforme all'interno del messaggio ricevuto, per evitare che all'interno del singolo blocco non si abbia, statisticamente, un numero di errori superiore a quelli correggibili.

$$\begin{bmatrix}1&0&0&1\end{bmatrix}\begin{bmatrix}1&0&0&0&0&1&1\\0&1&0&0&1&0&1\\0&0&1&0&1&1&0\\0&0&0&1&1&1&1\end{bmatrix}=\begin{bmatrix}1&0&0&1&1&0&0\end{bmatrix}$$

$$\begin{bmatrix}0&0&1&1&0&0&0\end{bmatrix}$$

$$\begin{bmatrix}1&0&1&0&1&0&0\end{bmatrix}\begin{bmatrix}0&1&1\\1&0&1\\1&1&0\\1&1&1\\1&0&0\\0&1&0\\0&0&1\end{bmatrix}=\begin{bmatrix}0&0&1\end{bmatrix}$$

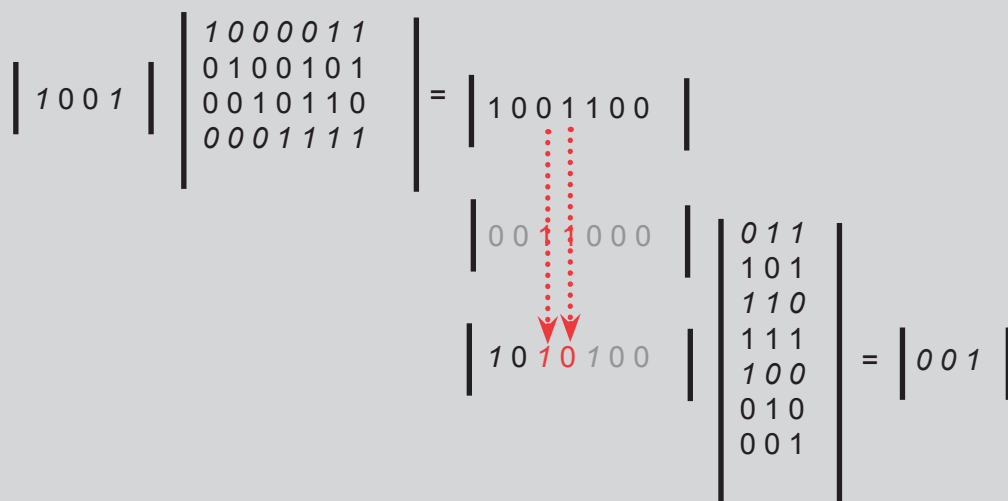



Fig. 5 - Esempio di superamento della capacità di correzione del codice.
In questo caso il canale introduce due errori, nell'esempio sul terzo e quarto bit.
La sindrome determinata in ricezione in questo caso fornisce una indicazione errata e il settimo bit viene considerato errato, il decodificatore quindi non corregge i due bit errati, ma introduce un ulteriore errore che fa si che la parola decodificata diventi [1 0 1 0 1 0 1].

6. Burst di errori e interleaving

Molto spesso gli errori vengono introdotti dal canale come salve o raffiche di errori (*error burst*). Ad esempio, nel caso della registrazione magnetica su nastro, la causa di errori può essere la mancanza di ossido o la presenza di particelle di polvere e quindi è interessata un'intera area del supporto e i bit ad essa corrispondente. Nel caso della diffusione, l'errore può interessare un simbolo demodulato, a cui corrispondono in genere più bit, oppure una temporanea riduzione della potenza del segnale ricevuto (*fading*) può comportare la mancata demodulazione di un certo numero di simboli.

Per ottenere una distribuzione uniforme degli errori e quindi limitare il più possibile gli effetti dei burst utilizza l'*interleaving* (figura 6), ovvero si intercalano le parole costituenti un blocco sufficientemente ampio di dati.

Questa tecnica non richiede un aumento della ridondanza, ma, grazie ad opportune memorie in registrazione e riproduzione, che però introducono un ritardo proporzionale alla lunghezza del codice a blocco e del fattore di *interleaving*, permettono di sfruttare al meglio le capacità correttive del codice.

L'*interleaving* a blocchi descritto in figura 6 è utilizzato in applicazioni dove è di scarsa importanza il ritardo introdotto per la scrittura nella memoria, quali i sistemi di videoregistrazione. D'altro canto una struttura regolare di intercalamento dei campioni può far si che gli eventuali errori non corretti creino una struttura anch'essa regolare di elementi di immagine errati, percepibile dal sistema psicovisivo. Per tale ragione in genere nel caso della videoregistrazione si utilizza una struttura non regolare nella memorizzazione dei campioni, in modo da diminuire la visibilità degli errori residui. Questa tecnica di *interleaving* viene denominata *shuffling* (in inglese, l'azione di mescolare le carte di un mazzo).

Nel caso dei sistemi di diffusione, dove è opportuno minimizzare i ritardi, in particolare per le trasmissioni in diretta, si utilizza uno schema differente, denominato *cross-interleaving,* che verrà descritto successivamente, nella seconda parte dell'articolo.

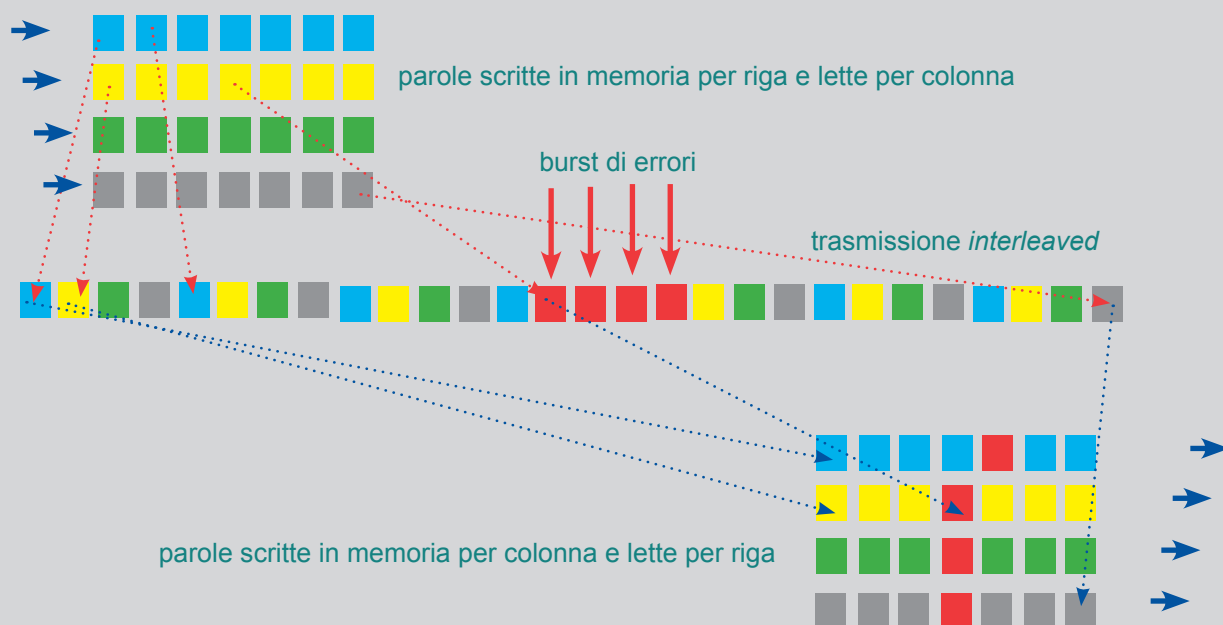



Fig. 6 - Esempio di interleaving: 4 parole di codice da 7 bit, ad esempio il codice di Hamming (7,4) in grado di correggere l'errore singolo, sono memorizzate e successivamente trasmesse sequenzialmente intercalando i bit relativi alle singole parole in modo che bit della stessa parola non siano consecutivi nel flusso trasmesso. Un eventuale burst di errori (in questo caso lungo 4) risulta così distribuito all'interno della memoria in ricezione: in questo esempio l'adozione dell'interleaving consente di recuperare tutta l'informazione, sfruttando appieno la capacità correttiva del codice. Ovviamente, sempre nell'esempio, se il burst fosse stato di lunghezza superiore, il codice non sarebbe stato in grado di correggere gli ulteriori errori.

7. Codifica Reed-Solomon

Un codice a blocco che ha trovato ampio impiego nel campo delle telecomunicazioni, ed in particolare per i sistemi per la trasmissione e diffusione TV digitale DVB e ATSC, è denominato RS (dal nome di Irving Reed e Gustave Solomon che lo proposero nel 1960 [7]). Questo codice ciclico è un caso particolare dei codici BCH.

Il codice non è in teoria utilizzabile nel caso di simboli binari, ma questa limitazione si supera nella pratica organizzando i bit sotto forma di parole costituite da un certo numero di bit, in questo modo si possono considerare tali parole binarie come i simboli su cui applicare il codice RS.

7.1 Il più semplice...

Ad esempio il codice RS può essere costruito utilizzando simboli costituiti da 3 bit, sono quindi disponibili 2^3 simboli differenti; la lunghezza di questo codice a blocco è pari a $n=2^3-1=7$ simboli.

Di questi 7 simboli, 5 sono i dati utili e 2 sono i simboli di ridondanza, che indichiamo con P e Q. Non si riporta qui la trattazione matematica che permette di calcolare, mediante un sistema di due equazioni, due sindromi denominate *locator* e *corrector* che, quando non sono nulle, sono utilizzate rispettivamente una per localizzare la posizione dell'errore singolo (del simbolo errato) e l'altra per correggerlo. Questo codice consente quindi di correggere fino a 3 bit errati, se sono parte dello stesso simbolo.

Un vantaggio dei codici RS è la possibilità di dimezzare la ridondanza richiesta se, nel caso in cui non serva la funzione "localizzazione" perché è già nota la posizione degli errori, si utilizza solo la funzione "correzione". Gli errori di cui è nota la posizione sono generalmente denominati *erasure* e questa tecnica trova applicazione nella registrazione su supporto magnetico o ottico i cui dati posso risultare "cancellati" (ad esempio per mancanza di ossido o per la presenza di particelle di polvere tra il traferro della testina e il nastro).

Ad esempio, il codice RS con simboli a 3 bit può correggere un simbolo errato all'interno del blocco da 7 simboli utilizzando due simboli di ridondanza, ma se è utilizzato nell'ambito di uno schema che preveda le segnalazioni delle *erasure* è sufficiente un solo simbolo di ridondanza (e quindi 6 sono quelli di informazione) per la correzione di quello di cui è già nota la posizione. In alternativa si possono utilizzare due simboli di ridondanza per correggere fino a due *erasure*.

7.2 ... e quello più usato

Al crescere del valore di *n* e di *k* aumenta il numero di equazioni che devono essere risolte dal decodificatore per localizzare gli errori. L'introduzione quindi di schemi di correzione così sofisticati, ma anche complessi, è stato quindi possibile solo ricorrendo a circuiti LSI. Molti degli standard di comunicazione attuali adottano i codici RS e quindi i produttori di DSP mettono oggi a disposizione librerie per consentire l'integrazione nei sistemi di codici RS con lunghezze con simboli da 3 a 12 bit, consentendo in questo ultimo caso di operare con blocchi fino a $2^{12}-1=4095$ simboli.

Quello più diffuso è quello che adotta simboli costituiti da byte, organizzazione tipica dei dati in campo informatico, cioè da simboli che possono assumere 2^8 valori differenti. Scegliendo simboli di 8 bit, la lunghezza ottimale del blocco è $n= 2^8-1=255$ simboli.

In funzione dell'applicazione, delle caratteristiche del canale e delle strategie di protezione si può scegliere il numero più opportuno di coppie di simboli utilizzati come *locator* e *corrector* o assegnare l'intera ridondanza alla correzione delle *erasure*.

La dimensione del blocco, 255 byte, è tale da rendere il codice particolarmente efficace nel caso in cui il canale sia caratterizzato dalla presenza di *burst* di errori molto lunghi. Il codice RS(255,239), in cui sono 16 i byte di ridondanza, fu utilizzato nella sperimentazione di trasmissione punto-punto via satellite in HDTV effettuata in occasione dei campionati mondiali di calcio Italia '90 e fu in seguito adottato nello standard ITU-T J81, per le reti di contributo televisive.

Il codice è in grado di correggere fino a $(n-k)/2$ simboli errati: e con $k=239$ si hanno 16 simboli di ridondanza, sufficienti a localizzare e correggere fino a 8 byte errati nel blocco, mentre se lo schema di codifica prevede che la localizzazione sia fatta a monte, la capacità di correzione sale a 16 *erasure*.

E' pertanto possibile correggere anche burst molto lunghi, fino a 8 o 16 byte.

Questo codice, in una versione accorciata, è utilizzata nei sistemi di diffusione televisiva DVB di prima generazione (DVB-S, DVB-C, DVB-T).

7.3 Codici accorciati

Si è visto che il parametro *n*, cioè la lunghezza dei blocchi, dipende dal numero di bit associati a ciascun simbolo. D'altro canto a volte è conveniente organizzare il flusso binario in funzione delle caratteristiche del messaggio emesso dalla sorgente o della struttura di sincronizzazione del multiplex e ciò restringe la scelta dei valori di *k* e *n*.

Una tecnica che consente di assegnare un numero prefissato *k* di simboli di informazione a ciascun blocco è quella di accorciare il codice prescelto (figura 7), anche se ciò avviene a spese di una diminuizione del tasso di informazione, cioè il rapporto *k/n*, che caratterizza il codice.

I sistemi DVB utilizzano il codice accorciato RS(204,188)[Nota 2] ricavato da RS(255,239) limitando a 204 il numero di byte utili trasmessi in ogni pacchetto, mentre il numero di byte di ridondanza rimane 16, mantenendo la capacità di correggere 8 byte oppure 16 nel caso di *erasure*.

8. Dai codici prodotto al picket code

I codici RS trovano impiego nei sistemi basati su nastro magnetico o disco ottico per la protezione dagli errori a *burst* che, a causa di imperfezioni del mezzo di registrazione o della presenza di materiale estraneo tra esso e il sistema di lettura, possono essere presenti nel flusso di dati riprodotti.

Normalmente tali sistemi usano codici prodotto ottenuti con una coppia di codici organizzati nello schema *outer* e *inner code* (figura 8).

Si tratta di codici accorciati derivati da Reed Solomon con *n*=255 e simboli costituiti da byte. Nel corso degli anni, al crescere delle capacità di elabo-

Nota 2 - La scelta di *k*=188 fu determinata dal requisito, ritenuto importante agli inizi degli anni '90, di facilitare il trasferimento delle informazioni audio e video compresse sui sistemi di ponti radio e di satelliti previsti per le telecomunicazioni, basati sullo standard ATM. Tale sistema, progettato per la commutazione e trasmissione telefonica , prevede che i dati siano organizzati in celle di 53 byte, di cui 48 byte di dati e 5 byte di intestazione. Il valore 188 (di cui i primi 4 costituiscono l'intestazione) fu scelto per facilitare la rimappatura dei dati utili codificati secondo gli standard MPEG nelle celle ATM.

Fig. 7 - Esempio di codice accorciato: si suppone di utilizzare un codice RS i cui simboli siano associati a 3 bit, in questo caso $n=2^3-1=7$ e con k=5 si è in grado di localizzare e correggere 1 simbolo errato. Il codice RS(7,5) implica quindi un blocco di 21 bit di cui 15 utili e 6 di ridondanza. Si supponga di voler organizzare i blocchi in modo che i bit utili siano solo 12. In questo caso è sufficiente calcolare i due simboli di ridondanza a partire da un simbolo fittizio, costituito da tutti 0, e da 4 simboli utili: ovviamente il simbolo costituito da tutti 0 non viene trasmesso e il blocco effettivamente posto sul canale è costituito da 12 bit utili e 6 di ridondanza: si parla di codice accorciato RS(6,4). In ricezione si provvede a ricostruire il blocco RS(7,5), prima di identificare e correggere gli errori.
Può succedere che il superamento della capacità correttiva del codice porti a localizzare errori anche nei simboli fittizi, ma in questo caso il decodificatore non effettua ovviamente alcuna correzione, ma rivela il superamento della capacità correttiva del codice e la presenza nel blocco di errori non correggibili, perché non localizzabili, agli eventuali stadi di correzione o mascheramento successivi.

razione dei decodificatori, è stato possibile adottare anche nel campo della registrazione codici con valori più elevati di *k* e *n*, approssimandosi a 255, e migliorando conseguentemente le prestazioni del sistema ECC, pur riducendo contemporaneamente la ridondanza ad esso associata.

Nel 1982 fu definito il formato del CD che adottava come sistema ECC la coppia di codici RS(32,28) e RS(28,24). Il primo, il codice interno, è in grado di correggere fino a due simboli e di rivelare la presenza di un numero superiore di errori. Il fattore di *interleaving* utilizzato è pari a 28. Il codice esterno è in grado di correggere fino a 4 *erasure*, garantendo così l'elevata capacità di correggere *burst* molto lunghi, fino ad un massimo di 4000 bit, corrispondenti a circa 2,5 mm in termini di lunghezza della traccia del disco.

Nel 1987 fu normalizzato il D1, il primo registratore digitale per segnale televisivo, a componenti e per usi professionali. Adottava un nastro da 3/4" e l'ECC era basata sulla coppia RS(32,30) e RS(64,60).

Nel 1991, il D3, sistema di registrazione televisiva digitale in formato composito su nastro da 1/2", fu in grado di migliorare notevolmente la densità di registrazione sul nastro, anche grazie all'adozione di un ECC basato sulla coppia RS(166,158) e RS(84,76).

Lo standard DVD, del 1995, utilizza una coppia

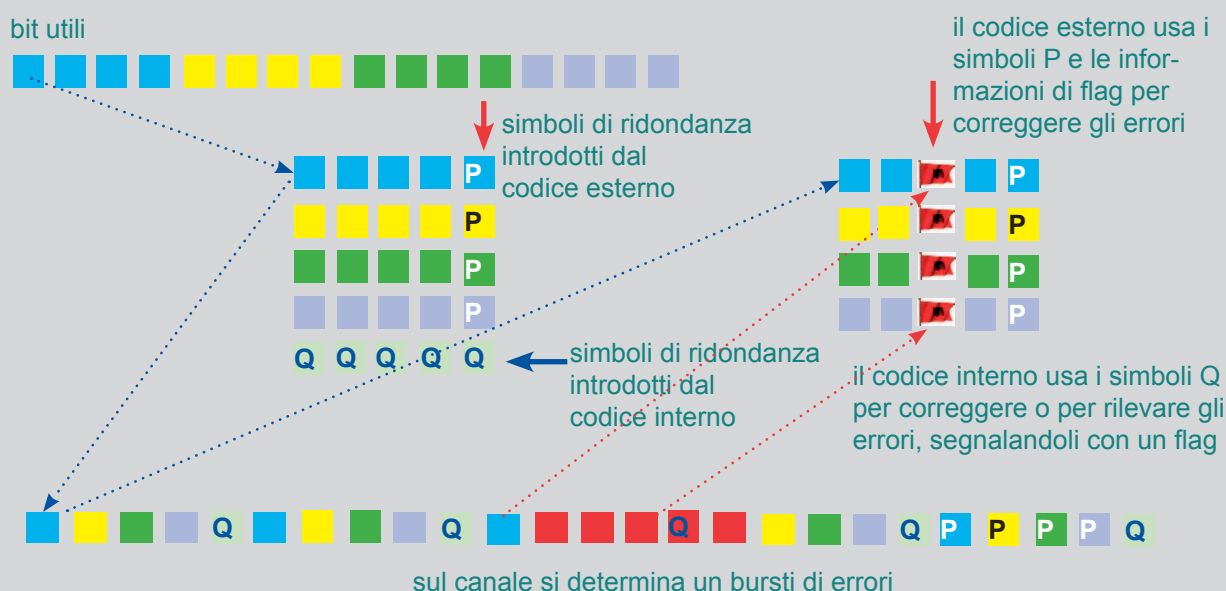



Fig. 8 - Esempio di schema di codice prodotto. Il flusso di dati utili è organizzato in blocchi, i simboli di protezione P sono calcolati in base ai dati disposti per righe; sucessivamente vengono calcolati i simboli Q, calcolati in base ai dati disposti per colonne.
Questo secondo codice è denominato codice interno (*inner code*) perché è applicato immediatamente prima del processo di trasmissione sul canale, mentre il codice che ha prodotto i simboli P è denominato codice esterno (*outer code*) perché è più lontano dal canale, nello schema codifica-canale-decodifica.
L'insieme dei dati vengono memorizzati o trasmessi sequenzialmente, sfruttando così la tecnica di *interleaving*. Nell'esempio si suppone che il canale introduca un burst di errori su cinque simboli consecutivi. In ricezione viene adottata la decodifica utilizzando i simboli Q che può consentire la localizzazione e correzione di errori. Nel caso in cui la capacità di correzione del codice interno sia superata, ma sia possibile rilevare la presenza di errori residui, non correggibili, questo evento è segnalato (con un *flag*) al secondo livello di correzione.
Nell'esempio si suppone che il codice interno sia in grado di correggere fino ad un simbolo errato e rivelare la presenza di errori in numero superiore, quindi corregge l'ultimo dei cinque simboli errati e rivela con un flag la presenza di errori nei quattro simboli di informazione che costituiscono la parola che precedente.
A questo punto il codice esterno, che si basa sui simboli ridondanti P, può essere in grado di correggere gli errori residui.

RS(182,172), come *inner code*, e RS(208,192), come *outer* code.

Nel 2000 è stato definito il formato di videoregistrazione digitale professionale a componenti e con compressione MPEG-2 su nastro da 1/2" denominato D10 (noto anche come Betacam IMX). L'ECC è basato sulla coppia RS(162,150) e RS(64,54): considerando anche le informazioni di sincronismo e di indentificazione dei blocchi, la ridondanza è pari al 28% e consente la corretta riproduzione dell'informazione video anche in caso di abrasioni pari 1,16 mm per le tracce video. I dati audio sono protetti con la coppia RS(137,125) e RS(18,8), ovvero il 147% di ridondanza per consentire la corretta riproduzione in caso di abrasioni pari 1,12 mm per le tracce audio.

Nel caso del disco ottico PDM, utilizzato per il sistema di videoregistrazione professionale XDCAM, è stato realizzato uno schema ECC denominato *Picket Code*. Tale schema è basato su una coppia di codici RS(62,30) e RS(248,216), denominati rispettivamente BIS e LDC. LDC è in grado di correggere *burst* di lunghezza molto elevata (tabella 1). I due codici sono intercalati in modo tale che vi sia un byte del codice BIS ogni 38 byte del codice LDC. Il codice BIS è molto robusto ed è in grado di correggere fino a 15 byte, che, grazie al fattore di *interleaving*, sono distribuiti su un blocco di dati lungo fisicamente 76,66 mm. La presenza di più errori consecutivi nei byte costituenti il BIS è utilizzato in fase di decodifica per rivelare la presenza e la posizione di eventuali burst di errori e ciò rende possibile sfruttare al meglio la capacità del codice LDC di correggere efficacemente le *erasure*, oltre agli errori isolati.

9. Un'incursione fra le schiere di dischi

Allo scopo di riassumere le tecniche di correzione degli errori precedentemente trattate, può essere utile analizzare una delle applicazioni più diffuse: la protezione delle informazioni memorizzate su hard-disk, per consentirne il recupero in caso di guasti.

Nel 1987, un gruppo di esperti di informatica dell'Università di California a Berkeley studiava metodi per realizzare schiere (*array*) di dischi (di basso costo, e quindi limitati in termini di capacità e tempi di accesso) che apparissero al computer ospite come un singolo disco, di grande capacità e con elevata velocità di lettura e scrittura.

Un primo metodo individuato consiste nel suddividere, a livello della scheda di controllo dei dischi, l'insieme di dati (considerato come una pagina) in tante strisce (*stripe*), ciascuna delle quali può quindi essere scritta (e letta) su uno degli *n* dischi che compongono la schiera. Poiché si può effettuare l'accesso in parallelo ad *n stripe,* il metodo permette di incrementare di un fattore prossimo a *n* la velocità accesso ai dati, ma il malfunzionamento o guasto di uno dei dischi può causare la perdita dell'intera informazione e la probabilità che ciò accada è proporzionale al numero di dischi, multipla della probabilità che si guasti un singolo disco.

	DVD	PDM
Distanza tra le tracce (Track Pitch) [µm]	0,74	0,32
Lunghezza di un Byte registrato [µm]	2,13	0,96
Area occupata da 1 Byte sul disco [mm^2/Byte]	1,576 10^{-6}	0,3072 10^{-6}
Numero di Byte per mm^2	635 KB	3,255 MB
Numero di Byte al mm	469 Byte	1041 Byte
Burst di errori correggibili	2790 Byte	9982 Byte
Lunghezza del burst correggibile [mm]	5,95	9,58

Tab. 1 - Confronto tra la capacità correttiva per burst di errori nel caso di un disco DVD e PDM. Il PDM utilizza luce a frequenza più elevata (blu-violetta) rispetto al DVD (rossa), e protetto in un *cartridge* per ridurre il rischio di contaminazione della superficie da parte di particelle estranee.
I dati riportati ipotizzano che una particella di polvere di 1 mm^2 sia depositata sulla superficie dei due media a confronto. I dati coinvolti sono, per il PDM, circa 5 volte in più, a parità di superficie, e circa il doppio a parità di lunghezza della traccia.
E' stato quindi sviluppato un sistema ECC, denominato *Picket Code*, che si basa su una coppia di codici Reed Solomon, RS(62,30) e RS(248,216), particolarmente adatto a rilevare burst molto lunghi, ma che non sfrutta lo schema classico di *inner* e *outer code*.

Occorre quindi individuare delle tecniche per proteggere le informazioni e gli esperti individuarono 5 livelli di protezione che denominarono RAID, cioè schiera ridondante di dischi a basso costo.

Il livello precedentemente descritto, che non fornisce alcuna protezione, ma non richiede nessuna ridondanza, fu definito il livello RAID 0.

Nel primo livello (RAID 1) i dati presenti su ciascun disco sono scritti anche su un secondo disco che quindi contiene, specularmente, tutta l'informazione (*mirroring*). Nel caso in cui uno dei dischi si guasti, può essere sostituito con uno nuovo, su cui si copiano i dati presenti sul disco ancora funzionante. Ovviamente occorre utilizzare il doppio dei dischi strettamente necessari e come minimo occorre avere due dischi. Dal punto dei sistemi di codifica per la protezione dagli errori, questo è un codice con ripetizione (ciascun bit è ripetuto due volte).

Il secondo livello (RAID 2) è basato sul codice di Hamming (7,4). Le informazioni sono divise in 4 *stripe* e scritte su 4 dischi, mentre altri 3 dischi contengono i 3 bit di parità calcolati in base ai 4 corrispondenti bit di informazione. Questo livello non ha mai trovato applicazione pratica, a causa del costo legato alla ridondanza, 7 dischi anziché 4.

I successivi due livelli (RAID 3 e RAID 4) organizzano i dati rispettivamente in byte o in blocchi,e li distribuiscono su *k* dischi; per ciascuno dei k bit presenti su tali dischi, calcolano un bit di parità che viene memorizzato su un ulteriore disco. Il numero totale di dischi è quindi $n=k+1$ e il numero minimo di dischi necessari è 3, due per i bit di informazione ed uno per i bit di parità. L'organizzazione dei dati in blocchi (RAID 4) è ovviamente più flessibile e può consentire tempi di accesso migliori rispetto all'organizzazione in byte (RAID 3).

Il livello RAID 5 è quello attualmente più utilizzato. I dati sono organizzati in blocchi, come nel caso del RAID 4, ma i blocchi di dati e di parità sono distribuiti uniformemente sugli $k+1$ dischi, anziché concentrare i blocchi di parità su un solo disco (figura 9). Questa struttura implica la stessa ridondanza rispetto ai due livelli precedenti e garantisce la stessa protezione in caso di guasto di un disco costituente la schiera, ma distribuisce uniformemente le operazioni di scrittura e lettura fra gli *n* dischi, migliorando le prestazioni dell'intero sistema. L'operazione di scrittura richiede mediamente più accessi al disco rispetto a quella di lettura poiché in genere, dopo l'operazione di scrittura, il *controller* rilegge il dato al fine di verificare l'assenza di errori, ed in caso di non coincidenza fra il dato scritto e quello riletto, effettua ulteriori tentativi.

I tre livelli (RAID 3, 4 e 5) consentono il recupero dell'informazione originaria nel caso di malfunzionamento o guasto di uno dei dischi. In genere il semplice bit di parità non permette la correzione dell'errore, solamente la rivelazione, ma in questo caso invece l'errore può anche essere corretto perché è nota anche la sua posizione: si sa quale degli *n* dischi è guasto. E' quindi sufficiente sostituire al disco guasto un disco di scorta e il sistema provvede

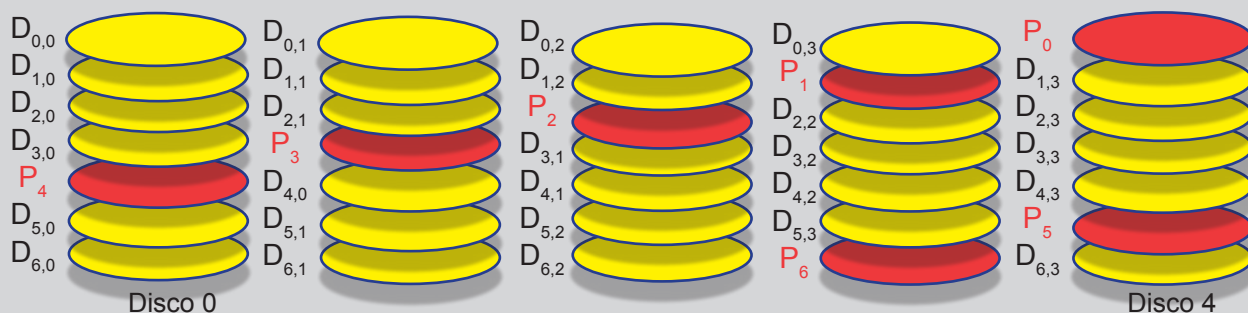



Fig. 9 - Semplice esempio di organizzazione dei dati nel caso del RAID 5, utilizzando 5 dischi: nella rappresentazione schematica in figura, ciascun blocco di dati è rappresentato da una ellisse, l'insieme di ellissi incolonnate rappresenta un disco.
Le parità P_0 ottenute a partire dai primi 4 blocchi di dati $D_{0,0}$...$D_{0,3}$ sono memorizzate nel quinto disco; P_1, ottenuto dai successivi 4 blocchi $D_{1,0}$...$D_{1,3}$ nel quarto disco ...

a rigenerare i dati perduti, calcolandoli a partire da quelli contenuti nei $k=n-1$ dischi funzionanti.

L'obiettivo originario, nel 1987, era l'uso di schiere di dischi di basso costo, ma il risultato è una struttura non necessariamente *Inexpensive*, cioè poco costosa, ma più affidabile, grazie alla possibilità di sostituire i singoli dischi: oggi il significato originario (*Inexpensive*) della I nell'acronimo RAID è stato modificato per evidenziare il fatto che i dischi sono indipendenti (*Independent*).

La protezione offerta dal livello RAID 5 è generalmente sufficiente, ma richiede che la sostituzione del disco guasto avvenga in tempi brevi, infatti il tempo che intercorre fra il verificarsi del guasto e la sostituzione è estremamente critico: un ulteriore guasto porta alla perdita di tutti i dati. Inoltre può essere a volte necessario disconnettere la schiera di dischi nel corso dell'operazione di sostituzione.

Al giorno d'oggi i server utilizzati per applicazioni video e audio, ad esempio per *streaming*, e in genere per web ed archivi, hanno dimensioni tali per cui richiedono schiere di dischi sempre più capaci. All'aumentare del numero di dischi che compongono il RAID aumentano sia la probabilità di guasto per il singolo disco che gli incovenienti causati dalla sospensione del servizio e pesanti sono i danni dovuti eventualmente al verificarsi di un guasto non recuperabile.

La probabilità di un difetto non recuperabile cresce con il numero di dischi che costituisce la schiera: nel corso della ricostruzione dei dati a seguito della sostituzione di un disco, il controller può rivelare un settore difettoso di uno dei dischi ritenuti funzionanti (evento con probabilità non trascurabile se i dischi sono numerosi) e i dati memorizzati in tale settore non possono più essere ricostruiti poiché manca la ridondanza, non più disponibile a causa della sostituzione del disco.

Sono stati pertanto sviluppati controller in grado di utilizzare un nuovo livello di protezione (RAID 6) basato sui codici MDS.

Tali codici sono un caso particolare del codice Reed Solomon con simboli costituiti da byte. Si suppone di utilizzare una schiera di 255 dischi, in modo da distribuire in dati sotto forma di parole di 255 simboli, di cui 253 sono le informazioni originali e 2 sono i simboli di ridondanza. Il codice è utilizzato per la caratteristica di poter correggere errori di cui è nota la posizione (*erasure*) e pertanto la possibilità di perdere informazione è limitata solo ai casi, estremamente poco probabili, in cui più di due dischi si guastino contemporaneamente, oppure se l'informazione "cancellata" debba essere ricostruita a partire da più settori che risultino difettosi al momento della sostituzione del disco.

Ovviamente non è necessario che la schiera RAID 6 sia costituita da 255 dischi: si può utilizzare una versione accorciata del codice (figura 10), basta garantire la presenza di due dischi aggiuntivi, per i simboli di ridondanza, cioè che sia $n=k+2$.

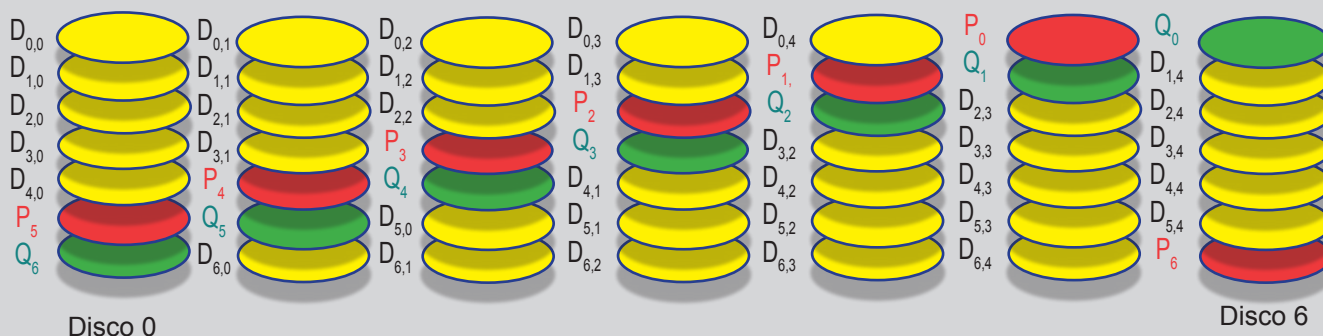



Fig. 10 - Semplice esempio di organizzazione dei dati nel caso del RAID 6, utilizzando 7 dischi (le colonne della tabella). Le parità P_0 e Q_0 ottenute a partire dai primi 5 blocchi di dati $D_{0,0}...D_{0,5}$ sono memorizzate rispettivamente nel sesto e settimo disco; P_1 e Q_1, ottenute dai successivi 5 blocchi $D_{1,0}...D_{1,4}$ nel quinto e sesto disco ...

Il vantaggio di utilizzare lo schema RAID 6, e quindi un codice più complesso come il Reed Solomon, rispetto a quello RAID 5, e quindi la semplice parità, è evidente nel caso in cui si considerino capacità complessive elevate ottenute con un grande numero di dischi e tenendo in considerazione la probabilità di trovare difetti latenti durante la fase di ricostruzione. Ad esempio con capacità complessive superiori a 5 TB ottenute con dischi da 320 GB, il passaggio da una schema RAID 5 ad uno RAID 6, può ridurre il tempo medio fra perdite di dati da valori dell'ordine di un mese a valori prossimi a 100 anni [8].

2 Prossimi al limite di Shannon

Testo e figure tratte da "Prossimi al limite di Shannon, 60 anni dopo"
di Marzio Barbero, Natasha Shpuza , Elettronica e Telecomunicazioni, Agosto 2008

1. Sessanta anni fa: un contributo leggendario

E' importante ricordare alcune delle tappe fondamentali che hanno consentito lo sviluppo delle teorie e delle tecniche che sono alla base dei sistemi di comunicazione moderni.

In particolare, basilare è il contributo costituito dalla coppia di articoli pubblicati proprio sessanta anni fa da Claude E. Shannon, destinata a rivoluzionare la tecnologia alla base delle telecomunicazioni [9].

Questo capitolo è strettamente correlato con tale panoramica, approfondendo maggiormente l'aspetto storico dei progressi teorici che hanno consentito il rapido sviluppo a cui assistiamo oggi.

L'evoluzione delle teoria dei codici per la protezione degli errori ha trovato dapprima applicazione soprattutto nel campo delle telecomunicazioni spaziali, ma negli ultimi anni sono state proprio le applicazioni di uso più generalizzato (diffusione televisiva digitale, telefonia mobile e *wireless*) a mettere in pratica in modo sempre più efficiente le teorie, anche quelle già pubblicate decenni fa, grazie alle possibilità offerte dai progressi dei circuiti VLSI in termini di velocità e di capacità di memoria.

In particolare il sistema di diffusione televisiva da satellite di seconda generazione (DVB-S2) è stato il primo a consentire di avvicinarsi come non era mai avvenuto prima al limite teorico della capacità di un canale digitale, quello indicato da Shannon sessant'anni fa.

2. 1948: il limite di Shannon

Scriveva Shannon: *"Il problema fondamentale della comunicazione è quello di riprodurre in un punto o esattamente, o approssimativamente, un messaggio definito in un altro punto."*

Shannon formulò una relazione fondamentale che consente di valutare la capacità C di un canale soggetto a rumore additivo con distribuzione gaussiana e caratterizzata da una larghezza di banda W.

"Può sembrare sorprendente che si debba definire un capacità C definita per un canale rumoroso, poiché non possiamo mai inviare informazione certa in un tale caso. E' chiaro, tuttavia, che inviando l'informazione in forma ridondante, la probabilità di errori può essere ridotta. ... Di fatto la capacità C precedentemente definita ha un significato completamente determinato. E' possibile inviare informazione alla velocità C attraverso il canale ***con una frequenza di errori o imprecisioni piccola a piacere*** *per mezzo di*

una codifica appropriata. Questa affermazione non è valida per velocità di segnalazione superiori a C."

La relazione fra capacità e larghezza di banda è fornita da:

"Teorema 17: La capacità di un canale di banda W perturbato da un rumore termico bianco di potenza N quando la potenza media del trasmettitore è limitata a P è data da:

$$C = W \log_2 \frac{P+N}{N}$$

La capacità C è espressa in bit al secondo per un canale soggetto a rumore additivo con distribuzione gaussiana (AWGN), W è la larghezza di banda del canale in Hertz, P e N sono rispettivamente le potenza del segnale trasmesso e la potenza di rumore espressi in Watt.

Sessanta anni fa veniva quindi definito il limite teorico della capacità di un canale binario e suggerito l'uso di *codici efficienti*, come quello di Hamming, per avvicinarsi a tale limite.

3. 1949-1962: L'evoluzione nella teoria dei codici

Il matematico Richard Hamming era stato assunto nel 1946 dai Bell Labs per lavorare sulla teoria dell'elasticità e utilizzava i computer del tempo, poco affidabili: nel caso in cui veniva rivelata la presenza di un errore l'esecuzione del programma si arrestava.

Hamming cercò una soluzione: organizzare i bit in blocchi, a cui aggiungere dei bit di parità in grado non solo di rivelare la presenza di un errore, ma anche di correggerlo, in modo da consentire ai programmi di completare i calcoli e giungere alla conclusione.

Nacque così il primo codice correttore di errori, H(7,4,3)[Nota 1] e Shannon lo descrive nel 1948 (figura 1) come *"un codice efficiente, che consente la correzione completa di errori e la trasmissione alla velocità C (fondato su un metodo dovuto a R. Hamming)"*.

I due paragrafi di descrizione del codice di Hamming contenuti nell'articolo di Shannon furono lo stimolo per l'articolo del 1949 di Marcel Golay [10].

Fig. 1

17. An Example of Efficient Coding

The following example, although somewhat unrealistic, is a case in which exact matching to a noisy channel is possible. There are two channel symbols, 0 and 1, and the noise affects them in blocks of seven symbols. A block of seven is either transmitted without error, or exactly one symbol of the seven is incorrect. These eight possibilities are equally likely. We have

$$\begin{aligned} C &= \text{Max}\,[H(y) - H_x(y)] \\ &= \tfrac{1}{7}\,[7 + \tfrac{8}{8}\log \tfrac{1}{8}] \\ &= \tfrac{4}{7}\ \text{bits/symbol}. \end{aligned}$$

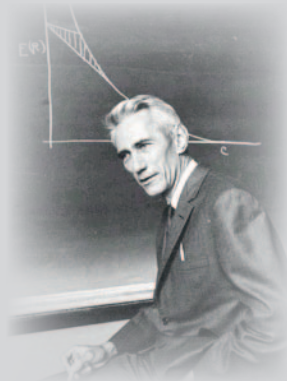
An efficient code, allowing complete correction of errors and transmitting at the rate C, is the following (found by a method due to R. Hamming):

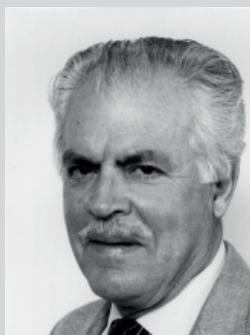
Let a block of seven symbols be $X_1, X_2, \dots X_7$. Of these X_3, X_5, X_6 and X_7 are message symbols and chosen arbitrarily by the source. The other three are redundant and calculated as follows:

X_4 is chosen to make $\alpha = X_4 + X_5 + X_6 + X_7$ even
X_2 " " " " $\beta = X_2 + X_3 + X_6 + X_7$ "
X_1 " " " " $\gamma = X_1 + X_3 + X_5 + X_7$ "

When a block of seven is received α, β and γ are calculated and if even called zero, if odd called one. The binary number $\alpha\,\beta\,\gamma$ then gives the subscript of the X_i that is incorrect (if 0 there was no error).

Nota 1 - Nella notazione (n,k,d_{min}) n indica la lunghezza del codice a blocco, k il numero di simboli che costituiscono il codice e d_{min}, la distanza minima di Hamming, è legata al numero di errori che il codice consente di correggere.

Ritratto di alcuni dei protagonisti

Claude Elwood Shannon: nato a Petoskey (Michigan) nel 1916 e morto nel 2001. Da ragazzo lavorò come telegrafista e conseguì due lauree nel 1936, in matematica e in ingegneria elettrica. Del 1940 è il dottorato, dal '41 al '72 lavorò ai Bell Labs. Durante la seconda guerra mondiale si occupò di ricerca per la guida di "missili". Dal '58 al '78 fu professore al MIT.

Richard Hamming: nato a Chicago nel 1915 e morto a Monterey (CA) nel 1998. Nel 1945 entrò a far parte del progetto Manhattan, il progetto per realizzare la bomba atomica, a Los Alamos. Nel 1946 iniziò l'attività nei Bell Laboratories dove restò fino al 1976.

Irving S. Reed, nato a Seattle nel 1923. ottiene il dottorato al California Institute of Technology nel 1949. Dal 1951 al 1960 associato al MIT. Professore, dal 1963, alla University of South California.

Andrew James Viterbi, nasce nel 1935 a Bergamo ed emigra con i genitori, a causa delle leggi razziali, nel 1939 negli Stati Uniti, dove studia al MIT dal 1952 al 1957. Ha fondato diverse società fra cui la Qualcomm, nel 1985.

Robert G. Gallager, nato a Filadelfia nel 1931, fa parte del personale tecnico dei Bell Labs nel 1953-54, ottiene il dottorato al MIT nel 1960. Professore e autore di numerosi testi sulla teoria dell'informazione.

Tale articolo è ritenuto da molti il più notevole mai scritto sulla teoria dei codici, perché in meno di una pagina vengono presentati: due codici *perfetti*[Nota 2], uno binario (23, 12, 7) ed uno ternario (11, 6, 5), la generalizzazione dei codici di Hamming e la prima pubblicazione di una matrice di controllo della parità.

L'articolo in cui Hamming descrive il codice è del 1950 [6], cioè fu pubblicato due anni dopo quello di Shannon. La spiegazione di tale ritardo rispetto alla citazione di Shannon, fu fornita da Hamming stesso: *"il lavoro fu concluso in tre mesi, ma per ragioni brevettuali fu tenuto in sospeso per due anni"*.

Sia i codici di Hamming che quelli di Golay sono *lineari*, cioè la somma modulo-q di una coppia di parole di codice costituite da simboli q-ari (cioè binari, ternari...) è anch'essa una parola di codice.

Nel 1954 Muller [11] descrive l'applicazione di codici nel contesto della progettazione in logica booleana, e Reed identifica tali codici come classe di codici lineari a blocco e ne propone l'algoritmo di decodifica [12].

Sempre nel 1954 viene descritto l'algoritmo di decodifica di Wagner, il primo algoritmo in letteratura di decodifica *soft-decision*. Questo è un approccio fondamentale per l'evoluzione della decodifica e per consentire le prestazioni oggi raggiunte: la decodifica tiene conto della affidabilità della decisione sul simbolo in uscita dal canale.

P. Elias inventa nel 1954 i codici prodotto [13] ed un anno dopo inventa i codici convoluzionali [14]. Il suo allievo, Robert Gallager, puntualizza che in tale articolo è evidenziato che *"l'ottenimento di una probabilità di errore piccola, a qualsiasi probabilità prossima alla capacità, richiede necessariamente un codice con una elevata lunghezza di blocco"*.

Successivamente (1957) sono scoperti i codici *ciclici* [15], che sono codici a blocco, lineari e che godono dell'ulteriore proprietà: lo shift ciclico di una parola di codice è ancora una parola di codice. Questa caratteristica consente di realizzare codificatori e decodificatori di limitata complessità. Inoltre tali codici possono essere descritti mediante un *polinomio generatore*. Sono anche denominati CRC; il loro uso è limitato generalmente alla sola rivelazione degli errori, infatti la complessità del decodificatore cresce esponenzialmente con il numero di errori correggibili.

Nota 2 - In termini matematici, in un codice perfetto, le sfere intorno alle parole di codice costituiscono una partizione dello spazio dei vettori.

Una sottoclasse dei codici ciclici è scoperta quasi contemporaneamente da Hocquenghem nel 1959 [16] e da Bose e Ray-Chaudhuri [17] nel 1960 e pertanto sono noti come codici BCH.

Sempre nel 1960 [7] Reed e Solomon descrivono i codici universalmente oggi noti con i loro nomi (RS): sono una classe non binaria dei codici BCH, o, in alternativa, i BCH sono sono sottocodici di un sottocampo di codici RS.

Nel 1962 [18], Gallager è motivato, nella sua tesi di dottorato con la supervisione di Elias, dalla ricerca di una classe di codici quasi casuali che possano consentire una decodifica prossima alla capacità del canale e caratterizzati da una complessità tale da non comprometterne la fattibilità: introduce così i codici LDPC. L'algoritmo di decodifica APP descritto è ritenuto la prima citazione in letteratura dell'algoritmo somma-prodotto oggi ampiamente utilizzato.

Nei primi quattordici anni, a partire dall'articolo di Shannon, erano state poste tutte le basi teoriche su cui si fondano gli sviluppi tecnologici realizzati anche in anni molto più recenti. Ma in quegli anni le applicazioni pratiche non erano state così rapide come l'evoluzione teorica.

4. Comunicazioni dallo Spazio

Un canale di comunicazione che ha ampiamente tratto vantaggio dall'uso delle tecniche di recupero delle informazioni in presenza di elevato rumore è quello delle comunicazioni spaziali.

Il 4 ottobre del 1957 l'Unione Sovietica lanciò nello spazio lo Sputnik, il 31 gennaio 1958 gli Stati Uniti, colti di sorpresa, risposero con il lancio di Explorer I, progettato e costruito da JPL (laboratorio fondato dall'Istituto di Tecnologia della California nel 1930) su richiesta dell'esercito americano. Era nata la competizione per lo spazio, che ebbe come conseguenza un ampio dibattito sul controllo civile o militare della spazio: il 29 luglio 1958, 50 anni fa, il presidente

Eisenhower firmava l'atto di nascita dell'agenzia civile NASA, di cui JPL è oggi Laboratorio.

Per consentire la comunicazione con le navicelle spaziali, in JPL furono messe a punto due tecniche fondamentali: l'uso di shift register per codificare le informazioni aggiungendo la ridondanza necessaria alla correzione degli errori e l'uso del PLL, indispensabile per agganciare la frequenza dell'oscillatore locale per demodulare le informazioni ricevute.

I motivi per cui sono state le comunicazioni verso lo spazio a dare inizialmente il maggior spunto agli studi sugli schemi di codifica di canale sono:

- il canale è affetto solo da rumore gaussiano bianco
- la banda è, di fatto, illimitata
- guadagni di frazioni di decibel hanno un valore economico e scientifico molto importante (la capacità di carico dei primi razzi era minima e la potenza disponibile a bordo per la trasmissione dei dati era bassa)
- la complessità, ed il corrispondente ingombro e costo, degli apparati di ricezione e di decodifica, può essere notevole

Nella missione Mariner del 1965 furono inviate con successo immagini di Marte da 200 x 200 pixel, ciascuno rappresentato da 6 bit (64 livelli) ad una velocità di 8 bit al secondo: la trasmissione di ogni singola immagine richiedeva circa 8 ore, non erano utilizzati codici.

Nelle missioni successive, dal 1969 al 1977, le cose migliorarono notevolmente con l'utilizzo dei codice Reed-Muller (RM). I codici RM sono caratterizzati da un insieme di parametri e la scelta flessibile dei valori ne ha consentito un ampio uso. Il guadagno di codifica offerto non era molto elevato, circa 3,2 dB, ma si stima che a quel tempo ogni dB di guadagno corrispondesse ad un risparmio sul costo della missione spaziale di circa un milione di dollari. Grazie al codice RM utilizzato, cui ad ogni 6 bit di informazione sono associati 26 bit per la correzione degli errori, il bit-rate era cresciuto a 16 kbit/s. Con la missione Viking (1976) le immagini sono trasmesse a colori, come tre componenti separate, una per ciascun colore primario.

Il codice era stato messo a punto da Irving Reed, chiamato nel 1963 alla JPL in quanto aveva assunto notorietà per aver sviluppato, alla RAND Corporation, un computer della dimensione di una scrivania, dieci volte meno ingombrante dei suoi contemporanei. Era stato Andrew Viterbi, in JPL dal giugno 1957, a suggerire l'assunzione di Reed.

E proprio a Viterbi è associato l'algoritmo di decodifica noto con il suo nome (VA), introdotto nel 1967 [19], che ha consentito di realizzare decodificatori veloci per i codici convoluzionali.

Infatti i codici convoluzionali furono fra i primi ad essere utilizzati per le comunicazioni spaziali, poiché il codificatore è realizzabile con una struttura estremamente semplice, basata su alcuni flip-flop e porte logiche. Nella missione Pioneer 9 (1968) in ricezione si utilizzava un minicomputer a 16-bit con clock da 1 MHz, decisione soft basata su campioni quantizzati a tre bit e decodifica sequenziale con algoritmo di Fano. Il bit-rate del canale era 512 bit/s.

L'algoritmo di Viterbi fu presto riconosciuto come algoritmo di decodifica ottimo per la decodifica convoluzionale, la cui realizzazione pratica, ad esempio con una macchina a 64 stati, poteva consentire guadagni dell'ordine di 6 dB. In effetti il decoder a 64 stati realizzato dalla Linkabit, fondata, fra gli altri, da Viterbi nel 1968, era *un grande mostro che riempiva un rack*, ma in grado di operare a 2 Mbit/s. Già nel 1975 era possibile integrare l'algoritmo di Viterbi in un chip, rendendolo disponibile per una più ampia gamma di applicazioni nelle telecomunicazioni.

Nelle missioni Voyager del 1977 è introdotto lo schema costituito dal codice RS(255,223,33), in grado di correggere fino a 16 byte errati, come *inner code* concatenato con il codice convoluzionale a 64 stadi, R=1/2, L=7[Nota 3]. Tale schema diviene lo standard NASA.

Nota 3 - il code rate R è il rapporto k/n dove n è il numero di bit in uscita dal codificatore in corrispondenza di k bit di informazione in ingresso: i bit in uscita sono generati in funzione dei k bit in ingresso e dei precedenti L-1 blocchi di k bit, dove L è denominata constraint length.

continua a pag. 28

Comunicazioni dallo Spazio Profondo

Pioneer 9

La serie di sonde Pioneer fu progettata per valutare l'operatività dei veicoli spaziali, su orbita solare. Nelle prime missioni, in base a comandi inviati da terra, potevano essere selezionati cinque bit-rate: 512, 256, 64, 16, e 8 bit/s.

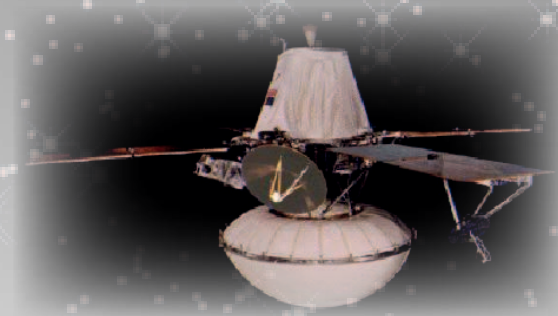
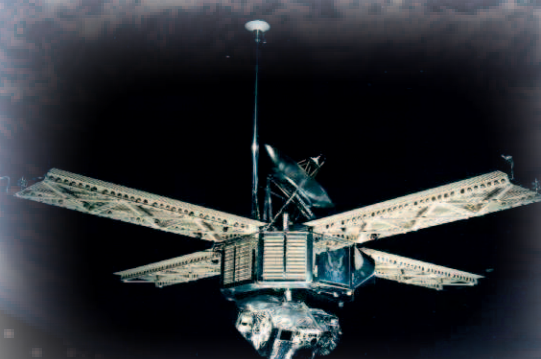
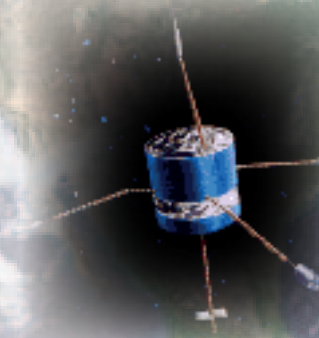
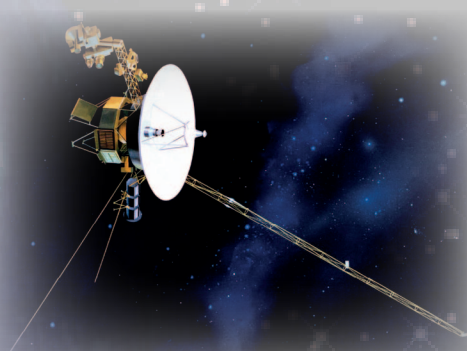
La sonda Pioneer 9, lanciata nel 1968, fu la prima ad utilizzare un codice convoluzionale, R=1/2, L=25, con decodifica sequenziale di Fano. Uno schema analogo, L=32, fu utilizzato nelle successive missioni Pioneer 10 (1972), 11 (1973) e 12 (1978) che esplorarono Giove, Saturno e Venere.

Mariner

Le sonde Mariner utilizzarono dal 1969 al 1973 un codice Reed-Muller RM(32,6,8) con R=0,1875, costituito da parole di 32 bit, 6 di informazione e 26 di parità, in grado di correggere fino a 7 errori. Il canale permetteva la trasmissione di 16 kbit/s verso la terra.

Il Mariner 6 (1969) registrò e inviò a terra 143 immagini della superficie di Marte. Il Mariner 9 (1971) fotografò il 100% della superficie di Marte e le sue lune Phobos e Deimos. Il Mariner 10, lanciato nel novembre 1973, nel febbraio '74 sorvolò Venere e nel marzo '74 Mercurio.

Viking

Analogo schema di codifica fu utilizzato per la missione Viking 1, lanciata nell'agosto 1975: la navicella in orbita fotografò la superficie di Marte e il modulo di atterraggio inviò foto a colori dal luogo dell'atterraggio. Dati pervennero fino al novembre 1982.

Voyager 1 e 2

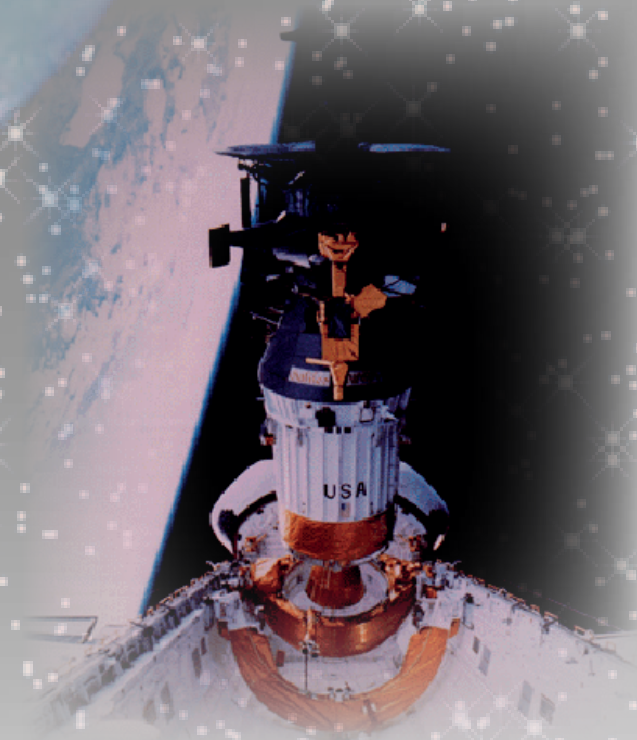
Le sonde Voyager 2 e 1 furono lanciate, rispettivamente, a maggio e settembre del 1977.

Lo schema di codifica utilizzato era basato su un codice convoluzionale, R=1/2 e L=7, con decodifica di Viterbi, concatenato con codice RS(255,233). Il decoder RS, grazie ad un hardware sviluppato appositamente, era in grado di operare a 1 Mbit/s. Il Voyager 2 ha utilizzato anche un codice di Golay.

La sonda Voyager 1 dal febbraio 1998 è diventato l'oggetto realizzato dall'Uomo più lontano dal Sole, avendo superato la distanza raggiunta da Pioneer 10.
Poiché ora la distanza è prossima alle 15 ore luce, i dati raccolti dalla sonda arrivano con tale ritardo al centro di controllo della JPL.
Voyager 2, lanciata prima della sua gemella, viaggia su un orbita meno veloce, si è "avvicinato" a Giove (1979), Saturno (1981), Urano (1986), Nettuno (1989) e ora lavora per la sua missione interstellare.
La NASA ritiene che il contatto potrebbe essere mantenuto con le due sonde oltre il 2020.

Galileo

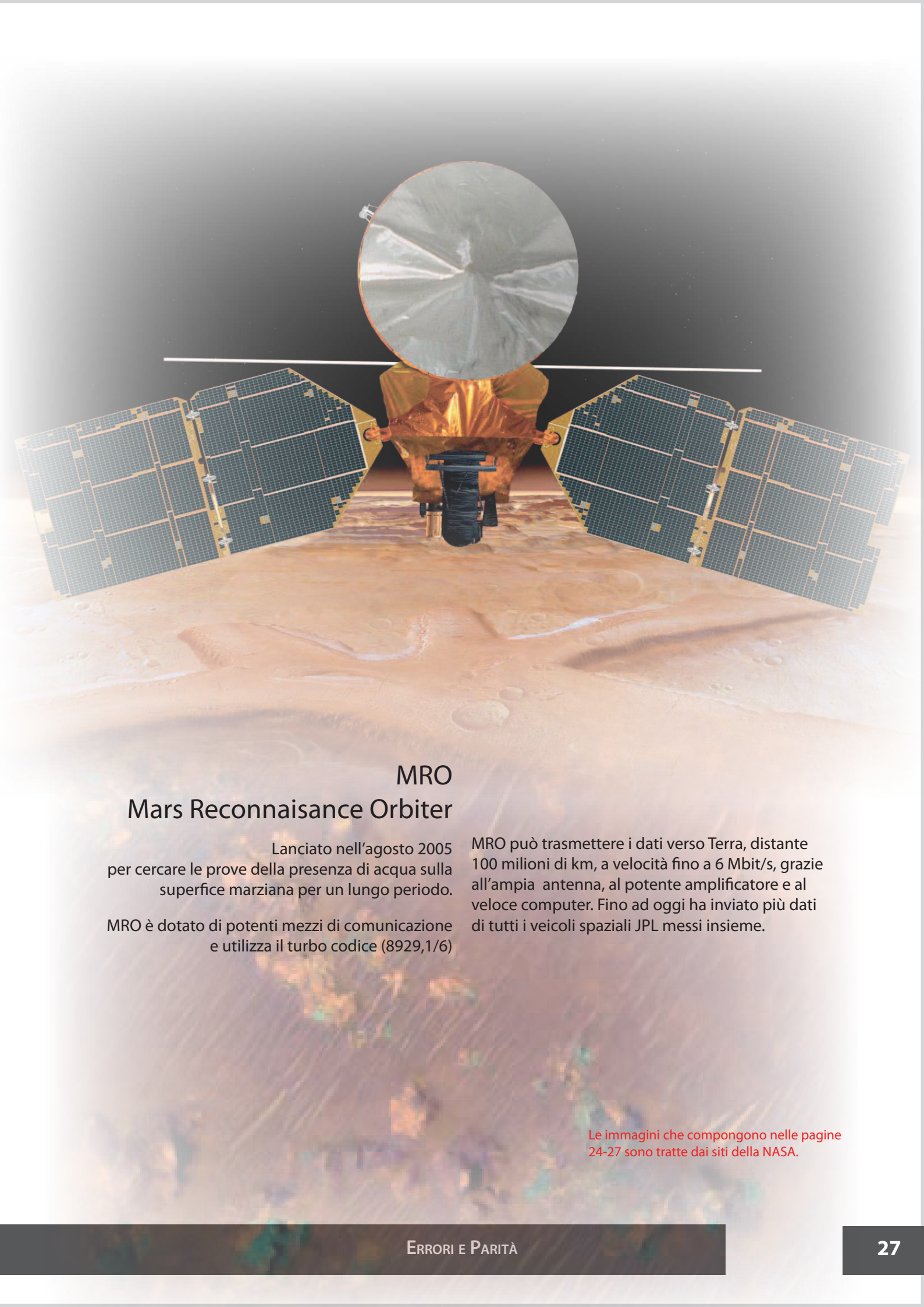
La missione era stata inizialmente progettata per inviare la navicella spaziale verso Giove con un viaggio diretto della durata prevista di circa tre anni e mezzo. Dopo l'incidente del Challenger (1986), per ragioni di sicurezza fu riprogettato il viaggio in modo da non richiedere l'uso di potenti stadi vettori. Il veicolo spaziale Galileo, portato in orbita dallo shuttle Atlantis, avrebbe sfruttato la forza gravitazionale interplanetaria, per raggiungere Giove in sei anni.

Atlantis decollò il 18 ottobre 1989, con Galileo nella stiva. Una volta iniziato il suo viaggio interplanetario, e a causa della modifica della traiettoria, la navicella spaziale Galileo fu soggetta a temperature molto più elevate di quelle originariamente previste. Il collegamento tra la sonda e terra era assicurato da due antenne a basso guadagno, mentre l'antenna ad alto guadagno, racchiusa come un ombrello, era protetta da scudi termici. Nell'aprile 1991, ormai sufficientemente lontano dal sole, l'antenna principale poteva spiegarsi per raggiungere il diametro previsto, di 4,8 m. Ma l'operazione non riuscì: l'ombrello non si aprì completamente.

Per consentire la trasmissione dei dati verso terra utilizzando le antenne a basso guadagno, dal 1993 al 1996 venne progettato uno schema di codifica molto più potente, in base al quale riprogrammare i codificatori di bordo, e che potesse consentire la decodifica al bit-rate previsto, grazie ad una struttura di elaborazione più complessa.

Lo schema di codifica concatenava un codice interno di tipo convoluzionale R=1/4, L=15, e un insieme di più codici esterni di tipo RS. La complessità hardware del decoder aumenta esponenzialmente al crescere di L, e con L=15 il numero di stati è pari a 2^{14}. Per la decodifica fu realizzato il BVD, basato su strutture di calcolo parallelo in grado di operare alle velocità richieste: ad esempio Galileo inviava dati telemetrici a 134,4 kbit/s.

Galileo arrivò in prossimità di Giove nel dicembre 1995, ed ha completato la sua missione il 21 settembre 2003, lanciato deliberatamente attraverso l'atmosfera gioviana.

Cassini

Lanciata nel 1997, è una missione congiunta della NASA, della ESA e della ASI. Prevedeva l'invio di un veicolo spaziale in orbita a Saturno per consentire lo studio del sistema del pianeta e dei suoi anelli per un periodo di quattro anni.

Per acquisire l'energia gravitazionale necessaria a raggiungere la destinazione, è passato accanto a Terra, Giove e Venere (due volte).

Cassini è entrato nell'orbita di Saturno il 1 luglio 2004, nel gennaio 2005 la sonda Huygen è discesa sulla superficie di Titano. Dal 2008 è iniziata una estensione di due anni della missione.

Il veicolo spaziale comunica con due antenne a basso guadagno e un'antenna a grande guadagno.

Lo schema di codifica adottato è basato su codice convoluzionale R=1/6, L=15 concatenato con un fattore I=5 al codice RS(255,233).

MRO
Mars Reconnaisance Orbiter

Lanciato nell'agosto 2005 per cercare le prove della presenza di acqua sulla superfice marziana per un lungo periodo.

MRO è dotato di potenti mezzi di comunicazione e utilizza il turbo codice (8929,1/6)

MRO può trasmettere i dati verso Terra, distante 100 milioni di km, a velocità fino a 6 Mbit/s, grazie all'ampia antenna, al potente amplificatore e al veloce computer. Fino ad oggi ha inviato più dati di tutti i veicoli spaziali JPL messi insieme.

Le immagini che compongono nelle pagine 24-27 sono tratte dai siti della NASA.

Nel 1993, a causa degli inconvenienti a dispiegare l'antenna principale nel corso della missione Galileo (vedere riquadro "Comunicazioni dallo Spazio Profondo") fu realizzato uno schema di decodifica a 2^{14} stati denominato BVD, in grado di operare ad una probabilità di errore dell'ordine di $2\ 10^{-7}$ con E_b/N_0 [Nota 4] di ≈ 0,8 dB e un guadagno di codifica reale di ≈ 10,2 dB.

5. 1993: I CODICI METTONO IL TURBO

Proprio nell'anno in cui lo schema basato su codice convoluzionale e codici RS raggiungeva il suo massimo (BVD), si verifica un evento [20] che dà un significativo impulso nella realizzazione di schemi di codifica che approssimano il limite di Sannon: i turbo-codici.

Nota 4 - E_0/N_0 è il rapporto fra E_b, l'energia media per bit, e N_0, la densità spettrale del rumore, cioè la potenza su una banda di 1 Hz.

Claude Berrou, un fisico professore di progettazione VLSI, interessato alla integrazione dell'algoritmo di Viterbi con decodifica iterativa, ipotizza che sia possibile migliorare la decodifica utilizzando un sistema a retroazione, con decodifiche ripetitive. Da qui il nome assegnato a questi codici: il decodificatore ottiene risultati sorprendenti utilizzando una retroazione (*feedback*), così come il motore turbo migliora le prestazioni riutilizzando parte dei gas di scarico.

I risultati delle simulazioni riportati dagli autori indicano come il limite di Shannon possa essere approssimato, con una differenza inferiore a 0,7 dB. Inizialmente tali risultati sono considerati con scetticismo, ma presto altri ricercatori ottengono risultati simili, e verificano che le prestazioni dei turbo codici dipendono dalla dimensione del codice *n* e dal fattore di *interleaving*: come aveva già indicato Elias nel 1954, le prestazioni del codice crescono al crescere della dimensione del blocco.

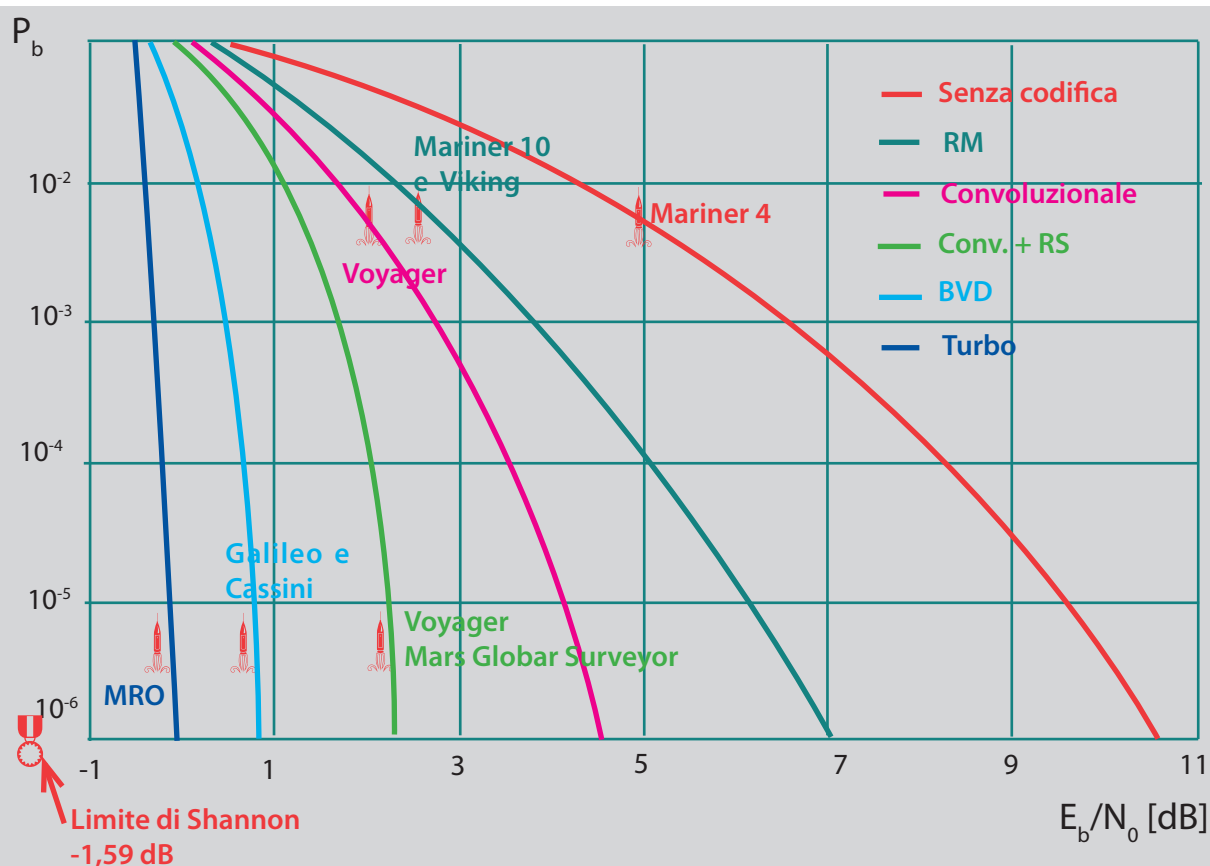



Fig. 2 - Le sonde Mariner 2,4, 5 non utilizzarono codifica per la protezione dagli errori. Mariner 6,7,9 10 e Viking utilizzarono il codice RM(32,6). Voyager 1 e 2, Magellan, Mars Global Surveyor utilizzarono il codice convoluzionale R=1/2, L=7 con decodifica di Viterbi. Galileo, Mars Pathfinder, Cassini e Mars Exploration Rover utilizzano il codice convoluzionale con L=15 e Big Viterbi Decoding. Messenger to Mercury e Mars Reconnaisance Orbiter utilizzano il turbo codice (8920, 1/6).

La prima missione spaziale a utilizzare il nuovo schema raccomandato dalla CCSDS basato su turbo codice è del 2005, il MRO (figura 2).

6. Codici e diffusione delle informazioni televisive

Ovviamente non sono solo le missioni spaziali a trarre vantaggio dai progressi nella codifica per la protezione dagli errori.

Nell'ambito televisivo e multimediale possiamo ricordare che il codice di Hamming è utilizzato nei servizi teletext (in Italia Televideo) introdotti negli anni '70. E' del 1982 la definizione del formato del CD, seguita da quella del DVD nel 1996: entrambi i formati utilizzano due codici RS, in uno schema di codifica a prodotto.

I sistemi digitali di diffusione televisiva da satellite sono stati resi possibile fin dall'origine dall'impiego di schemi sofisticati di protezione dagli errori. Infatti i sistemi di compressione delle informazioni video e audio, riducendo al minimo la ridondanza e non prevedendo la ritrasmissione dei messaggi in caso di errori, richiedono che la probabilità di errore per il flusso di dati ricevuti sia molto bassa, dell'ordine di 10^{-9}.

Gia nel primo esperimento, in occasione dei campionati mondiali di Italia '90, di trasmissione del segnale in Alta Definizione si utilizzò uno schema di codifica basato su codice esterno RS(255,239) e interno convoluzionale. Il bit rate complessivo, di poco inferiore a 70 Mbit/s, era trasmesso utilizzando una coppia di trasponder del satellite Olympus, essendo la capacità massima dei modulatori e del canale costituito da ciascun trasponder di circa 34 Mbit/s.

Lo standard DVB-S (1996) adotta uno schema basato su codice esterno accorciato RS(204,108) seguito da interleaver con profondità 12 e codice interno convoluzionale R=1/2, L=7.

Gli standard DVB-RCS e DVB-RCT adottano turbo codici.

E il DVB-S2, lo standard di seconda generazione per la diffusione da satellite, adotta uno schema di protezione basato sui codici LDPC.

7. Oggi: il ritorno dei codici LDPC

Dopo il 1993, a seguito dell'avvento dei turbo codici, molti ricercatori riprendono in considerazione anche schemi di codifica le cui prestazioni non erano state nel passato considerate soddisfacenti, a causa dei vincoli in termini di complessità di calcolo e capacità di memoria. In particolare l'attenzione si focalizza sui codici proposti da Robert Gallager nel 1962 [18] e viene dimostrato che prestazioni molto prossime al limite di Shannon possono essere raggiunte con codici LDPC di grandi dimensioni e decodifica iterativa [21].

Lo standard DVB-S2 (2003) [22] adotta la concatenazione di due codici: un BCH come codice esterno e un LDPC come codice interno. Il codice LDPC ha una lunghezza di parola che può essere n=64800 (per le trame normali) o n=16200 (per le trame corte). Approssima il limite di Shannon entro 0,6÷0,8 dB ed è realizzabile grazie alle tecniche di integrazione attuali [23].

Il DVB-S2 è utilizzato per la diffusione via satellite di programmi televisivi ad Alta Definizione (HDTV).

E' uno degli elementi fondamentali per la dimostrazione di trasmissione via satellite delle immagini Super Hi-Vision (SHV), realizzata in occasione della IBC 2008 [24]. Le immagini SHV sono composte da un numero di pixel 16 volte superiore a quelle HDTV, grazie all'evoluzione dei sistemi di compressione video è possibile ridurre il bit-rate a 140 Mbit/s, che, poiché non sono attualmente disponibili demodulatori adatti a tale capacità, vengono suddivisi in due flussi da 70 Mbit/s.

Anche il sistema di diffusione terrestre di seconda generazione, il DVB-T2, adotta la codifica LDPC per la protezione dagli errori [25].

Altri standard che utilizzano questi codici sono le nuove versioni di WiMax mobile (IEEE 802.16e-2005) e WiFi (IEEE 802.11n).

8. Conclusione

Sembrerebbe quindi che i sistemi attuali raggiungano il limite indicato sessanta anni fa da Shannon e che sarà pertanto difficile assistere in futuro a progressi significativi in questo campo.

Molte delle informazioni utilizzate per la stesura di questo articolo sono tratte da [26][Nota 5], e in tale articolo si cita un Workshop, tenuto in Florida nell'aprile 1971, che è ricordato come "*la codifica è morta*", perché sembrava che nulla di nuovo potesse essere aggiunto a quanto fino allora pubblicato sul tema.

Abbiamo visto che l'avvento, del tutto imprevisto, dei turbo codici nel 1993 aprì nuove prospettive per consentire l'avvicinamento al limite di Shannon.

Da questo punto di vista può quindi risultare azzardata la conclusione in [23][Nota 6]: *"I codici LDPC di DVB-S2 approssimano il limite di Shannon a 0,6÷0,8 dB ... Può risultare difficile giustificare la loro sostituzione nei prossimi decenni a venire."*

Nota 5 - Per approfondimenti sui codici si rimanda al numero speciale "Turbo.Information Processing: Algorithms, Implementations & Applications", del giugno 2007 di Proceeding of the IEEE.
Ricchi di informazioni sono inoltre i siti delle organizzazioni ASI, ESA, NASA, CCSDS, le cui URL sono indicate nella tabella che riporta acronimi e sigle.

Nota 6 - Per quanto riguarda il DVB-S2, si rimanda a "Special Issue on The DVB-S2 Standard for Broadband Satellite Systems", guest editors Alberto Morello and Ulrich Reimers, Int. J. Satell. Commun. Network., vol. 22, No 3, May-June 2004.

3 I codici: convoluzionali, turbo e LDPC

Testo e figure tratte da "Rivelazione, Correzione e Mascheramento degli Errori - Parte II" di Marzio Barbero, Natasha Shpuza , Elettronica e Telecomunicazioni, Agosto 2010.

1. Introduzione

I primi due capitoli evidenziano il ruolo dei codici convoluzionali e codici RS (Reed Solomon) sia nelle comunicazioni spaziali, sia nei sistemi di diffusione televisiva di prima generazione (DVB-S e DVB-T). Il § 2 che segue è dedicata ad una illustrazione della codifica convoluzionale, avendo già trattato i codici RS nel primo capitolo, § 7.

Si è visto nel secondo capitolo § 7 che la riscoperta dei codici LDPC, proposti nel lontano 1962 da Robert Gallager, è alla base delle prestazioni dei sistemi utilizzati per le missioni spaziali del futuro e dei sistemi di diffusione televisiva di seconda generazione (DVB-S2, DVB-T2 e DVB-C2). Le prestazioni di questi sistemi raggiungono il limite teorico di Shannon (secondo capitolo, § 2). A questi codici è dedicata il § 4 di questo capitolo.

2. Codifica convoluzionale

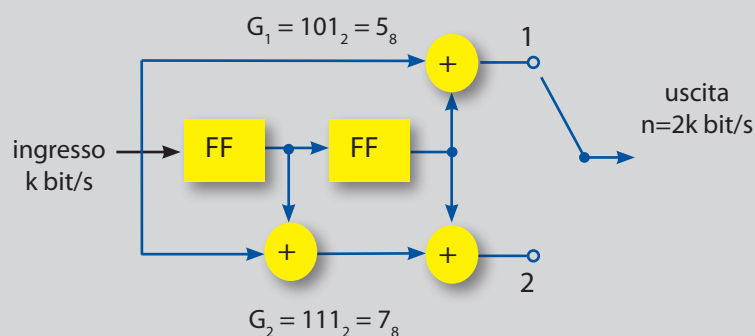
La codifica convoluzionale e la codifica a blocchi costituiscono le due forme principali di FEC. Essi differiscono fra loro poiché i codici convoluzionali non spezzano il flusso di dati da codificare in blocchi di lunghezza fissa, bensì la ridondanza è aggiunta in modo continuo al flusso codificato.

Si è visto nel capitolo precedente che la codifica convoluzionale ha trovato impiego sin dalle prime missioni spaziali, grazie alla semplicità costruttiva del codificatore, realizzabile grazie a pochi flip-flop e alcune porte logiche.

In figura 1 è rappresentato, come esempio, il "miglior" codice con *rate* R=1/2 e L=3.

Il *code rate* R e la *constraint length* L sono i due parametri principali che caratterizzano il codice. Nella codifica convoluzionale si opera sul flusso binario seriale: R è il rapporto k/n dove n è il numero di bit in uscita dal codificatore in corrispondenza di k bit di informazione in ingresso. I bit in uscita sono generati in funzione dei k bit in ingresso e dei precedenti L-1 blocchi di k bit, per cui si ha memoria del flusso di dati già codificato. Un ulteriore parametro che caratterizza il codice è la distanza libera d_{free} (*free distance*), la distanza di Hamming minima fra differenti sequenze codificate.

La struttura di un codificatore di questo tipo corrisponde a quella di una macchina a stati finiti. Il numero di stati cresce esponenzialmente con il crescere di L: per valori piccoli di L nella decodifica è utilizzato l'algoritmo di Viterbi (secondo capitolo, § 4), che presenta il vantaggio di richiedere un tempo fisso per la decodifica. In pratica si confronta una sequenza piuttosto lunga di bit ricevuti con tutte le

Fig. 1 - Il codificatore nello schema è composto da due componenti elementari: flip-flop e sommatori binari (XOR). E' caratterizzato da un *code rate* R=1/2, quindi il numero di bit in uscita è il doppio di quelli entranti. Il flusso in ingresso è applicato al registro a scorrimento che dispone di uscite intermedie (prese) in corrispondenza di ciascun stadio; ad ogni bit in arrivo all'ingresso, il flusso avanza di un colpo di clock e l'uscita è ottenuta prelevando alternativamente i bit ottenuti alle due uscite 1 e 2; il bit-rate in uscita è quindi doppio rispetto a quello in ingresso. La *constraint length* L corrisponde al numero di prese e la presenza o meno delle connessioni alle uscite dei flip-flop corrispondono ai polinomi generatori del codice G_1=101 (cioè manca la presa intermedia) e G_2=111 (l'uscita 2 dipende dai tre bit presenti nello shift). I polinomi generatori sono generalmente espressi in ottale e quindi questo codice è noto come L=3 (5,7), d_{free}=5.

possibili sequenze e si sceglie quella più prossima (criterio di massima verisimiglianza), ricavando da essa k bit ogni n bit ricevuti.

La decodifica introduce un ritardo, proporzionale alla sequenza esaminata, e una decisione errata può influenzare anche le successive decisioni: si può avere propagazione degli errori ed in tal caso gli errori si presentano a burst.

Per limitare gli effetti dei burst di errori si può ricorrere alla tecnica di due codici posti in cascata, un codice esterno seguito da un codice interno, così come si è visto parlando dei codici prodotto. In questo caso però si utilizza il termine codici concatenati (*concatenated codes*) per indicare questa tecnica.

Lo schema adottato per le comunicazioni spaziali a partire dal 1977 era basato su il codice convoluzionale R=1/2 e L=7 come codice interno, da simboli costituiti da byte protetti da un codice esterno RS(255,233,33), in grado di correggere fino a 16 byte errati.

Fig. 2 - Nel sistema DVB-S lo schema del trasmettitore prevede un codice esterno accorciato RS (204,188), seguito da un interleaver con profondità 12 e dal codice interno, convoluzionale con code rate R=1/2, *constraint length* L=7, polinomi generatori $G_1=171_8$ e $G_2=133_8$, d_{free}=10.

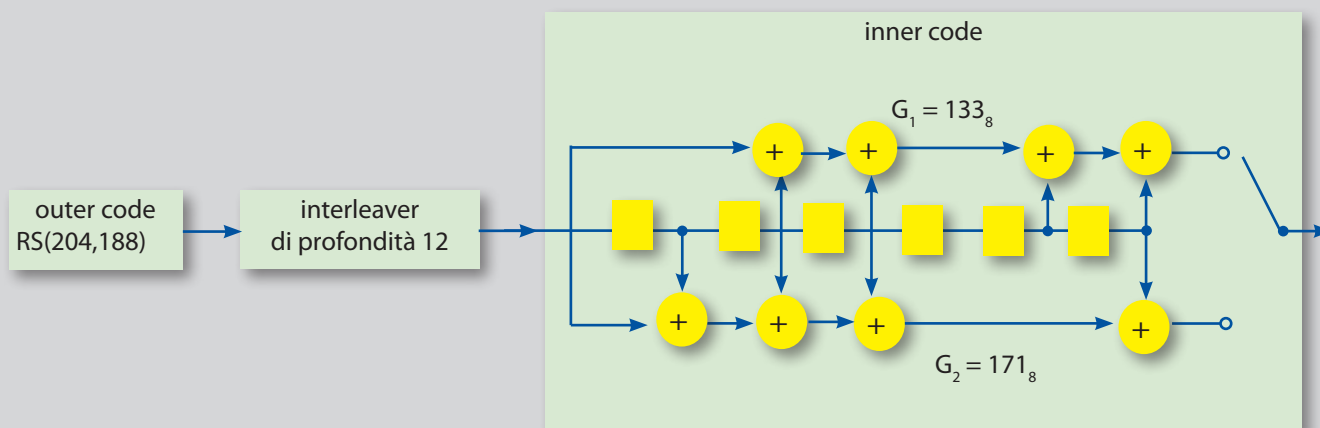
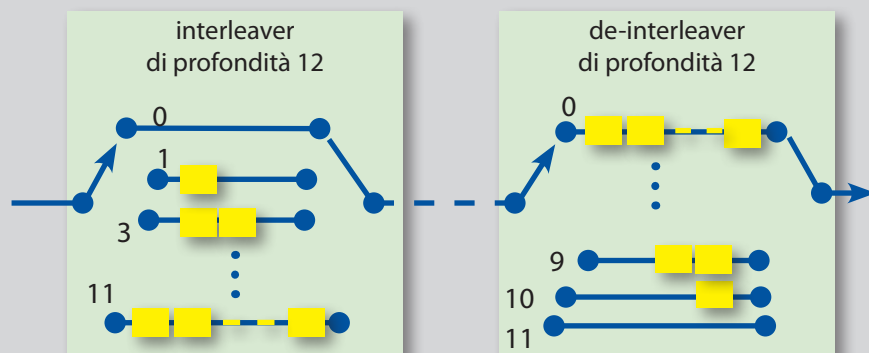



Fig. 3 - Nel sistema DVB-S l'interleaver consiste in 12 rami, ciclicamente interconnessi al flusso di byte in ingresso: ciascun ramo j è costituito da uno shift register (FIFO) di Mj celle. Il valore di M è pari a 17, rapporto fra il numero di byte costituenti il codice (204) e la profondità di interleaving (12). Al ramo 0 non corrisponde alcun registro a scorrimento, e quindi il ritardo introdotto è nullo, al ramo j=11 corrisponde un registro di 17*11=187 byte. Il de-interleaver ha una struttura analoga, ma l'indice j=0 corrisponde al ritardo massimo (187 byte) e quello 11 al ritardo nullo.

Anche nel caso del sistema di diffusione da satellite di prima generazione, il DVB-S, è stato adottato uno schema analogo (figura 2), basato su un codice convoluzionale *R*=1/2 e *L*=7. In questo caso si adotta un RS accorciato RS(204,188) derivato dal RS(255,239) e in grado di correggere fino a 8 byte errati o 16 byte nel caso di *erasure* (primo capitolo, § 7.3).

Per ridurre gli effetti dei burst di errori è utilizzata la tecnica dell'interleaving (primo capitolo, § 6). Lo schema di interleaving (usato in trasmissione) e de-interleaving (utilizzato in ricezione) convoluzionale è quello di figura 3.

La codifica convoluzionale permette la ricezione in caso di canali particolarmente rumorosi, ma richiede un'elevata ridondanza. Con *R*=1/2, metà della capacità del canale è utilizzata per la ridondanza.

Si può ridurre tale ridondanza, per sfruttare maggiormente la capacità per trasmettere i dati utili, se le caratteristiche del canale lo consentono e se si può accettare una riduzione di d_{free} e la conseguente minore robustezza.

A tale scopo si utilizza la tecnica di *puncturing*, cioè si "perfora" il flusso di dati avviati al modulatore, ovvero si eliminano alcuni dei bit in base ad una opportuna matrice (tabella 1). In fase di decodifica è noto il valore di code rate *R* utilizzato dal codificatore: tanto più è prossimo a 1, tanto più bassa è la ridondanza e meno robusto il codice. Il valore di *R* può essere variato nel tempo in funzione delle caratteristiche dal canale; si consideri ad esempio il caso di una sonda spaziale che si allontana dalla terra: al diminuire dalla potenza ricevuta, si accetta una riduzione della velocità dei dati ricevuti, pur di continuare l'acquisizione delle preziose informazioni raccolte.

3. I Turbo Codici

Nei primi anni '90 l'uso dei codici a blocco di tipo RS e di quelli convoluzionali distavano ancora da quelle teoricamente raggiungibili (limite di Shannon) di più di 3dB, ovvero gli schemi di codifica praticamente realizzabili allora richiedevano che i sistemi di comunicazione dovessero utilizzare, a parità di prestazioni, un'energia almeno doppia della minima teorica.

Come accennaot nel capitolo precedente, nel 1993 due ingegneri elettronici francesi, Claude Berrou e Alain Glavieux, proposero uno schema di codifica che consentiva un miglioramento tale da approssimarsi al limite di circa 0,7 dB [20].

continua a pag. 35

Tab. 1 - Matrici di perforazione usate per le telecomunicazioni satellitari e specificate per lo standard DVB-S. Se, ad esempio, si vuole utilizzare il codificatore di figura 1 con un *code rate R*=2/3, si invieranno solo i bit in posizione pari in uscita dal ramo superiore e tutti i bit in uscita dal ramo inferiore.

R	1/2	2/3	3/4	5/6	7/8
puncturing matrix	1 1	10 11	101 110	10101 11010	1000101 1111010
d_{free}	10	6	5	4	3

La Medaglia d'Onore IEEE a Viterbi

Montreal, 26 giugno 2010 - La prestigiosa IEEE Medal of Honor è stata consegnata a Andrew J. Viterbi per i suoi "contributi fondamentali alla tecnologia e alla teoria delle comunicazioni". Questo riconoscimento è, dal 1917, il più alto riconoscimento della IEEE, "la più ampia associazione a livello mondiale dedicata all'avanzamento dell'innovazione tecnologica e all'eccellenza a beneficio dell'umanità".

Andrew Viterbi è il secondo italiano a ricevere la Medaglia, il primo fu Guglielmo Marconi, nel 1920.

Andrea Viterbi nasce nel 1935 a Bergamo ed emigra nel 1939 con i genitori, a causa delle leggi razziali, negli Stati Uniti, dove il nome, al momento della naturalizzazione, diventa Andrew. Si laurea al MIT di Boston in ingegneria elettrica nel 1957. Dopo aver lavorato alla Raytheon, si trasferisce in California, al Jet Propulsion Laboratory, dove lavora per i sistemi di telemetria dei missili teleguidati. Nell'autunno del 1963, dopo aver acquisito il dottorato, insegna comunicazioni e teoria dell'informazione all'Università di Los Angeles.

Nel marzo 1966 semplifica l'algoritmo fino ad allora utilizzato per la decodifica dei codici convoluzionali: invece di operare successive iterazioni, il nuovo algoritmo, basato su una struttura a traliccio (*trellis*), considera i bit accanto a quello su cui di deve operare una decisione (è 0 oppure 1) ed effettua la scelta su base probabilistica. Il software opera una decisione sulla base di un numero limitato di stati, tipicamente da un minimo di 4 ad un massimo di 1000. L'articolo che riporta il nuovo algoritmo è pubblicato nel 1967 in *IEEE Transaction on Information Theory* [19].

Nel 1968 fonda la *Linkabit*, che, oltre a modem satellitari per applicazioni civili e militari, introduce uno scramber per la TV via cavo, utilizzato fino al 2008. Nel 1985 avvia una nuova impresa, la *Qualcomm* che si specializza nelle comunicazioni *spread-spectrum*, inizialmente per le comunicazioni da satellite e successivamente per la telefonia mobile.

Il CDMA (*Code-Division Multiple Access*) proposto dalla Qualcomm dal 1993 è parte degli standard per telefonia mobile 2G e 3G.

Lasciata la Qualcomm nel 2000, investe parte dei centinaia di milioni di dollari ricevuti come *stock option* nella società di investimento *Viterbi Group*, specializzata, ovviamente, in start-up nel campo delle comunicazioni.

Fra coloro hanno ricevuto la Medal of Honor, possiamo individuare, oltre a Viterbi, altri protagonisti della storia delle telecomunicazioni digitali e dei codici di protezione dagli errori, oggetto della serie di articoli che si conclude con quello pubblicato in questo numero: Harry Nyquist (1960), Claude E. Shannon (1966) e Robert G. Gallager (1990).

Nel traliccio, ciascun nodo corrisponde ad uno stato distinto ad uno specifico istante di tempo e ciascuna freccia rappresenta la transizione ad un nuovo stato successivo. L'algoritmo di Viterbi per la decodifica del flusso binario, codificato con codici convoluzionali, si basa sulla proprietà che il costo di un percorso lungo il traliccio può essere espresso come somma dei costi di transizione fra i nodi adiacenti nel tempo. Ad ogni passo è determinato il costo relativo alla transizione a ciascuno dei nodi successivi e solo i percorsi che hanno il costo minimo sopravvivono, gli altri sono eliminati.

Questi codici vennero denominati Turbo Codici perché nella decodifica si utilizza un percorso di retroazione, analogamente a quanto avviene in campo automobilistico con i motori turbo.

Un turbo codice è formato dalla concatenazione parallela di due codici separati da un interleaver.

In figura 4, a titolo di esempio, è riportato lo schema del codificatore turbo utilizzato per il sistema di telefonia mobile di terza generazione UMTS.

In generale la scelta dei codificatori e dell'interleaver è libera, ma la maggior parte delle realizzazioni si basa sui criteri adottati per quello in figura: i due codificatori sono identici; il codice è sistematico, ovvero i bit in ingresso sono anche presenti all'uscita; l'interleaver legge i bit in ordine pseudo-casuale.

La scelta dell'interleaver è fondamentale per il progetto dello schema di un turbo codice. L'interleaver pseudo casuale o pseudo-random è definito da un generatore di numeri pseudo casuali o da una look-up table.

Ha due scopi fondamentali.

Essendo posto all'ingresso del secondo encoder, la sua uscita ha caratteristiche statistiche completamente diverse dall'uscita del primo encoder, in particolare per quanto riguarda il peso, cioè il numero di 1 presenti.

L'uso dell'interleaving pseudo-casuale all'ingresso del secondo encoder rende i due flussi ottenuti completamente scorrelati, anche in uscita dai due decoder corrispondenti, e ciò è particolarmente vantaggioso in fase di decodifica.

Fondamentale per ottenere prestazioni ottimali è l'impiego di un metodo *soft -decision* per la decodifica. In un decoder *SISO* la decisione non è basata su una soglia *(hard-decision*) per decidere se il simbolo ricevuto è 0 oppure 1, ma il decoder elabora un valore reale (*soft*) ottenuto dal demodulatore e fornisce in uscita per ciascun bit una stima della probabilità che il bit trasmesso sia un 1.

Nel decoder turbo, le uscite dei due decoder forniscono stime degli stessi bit, ma i bit sono trasmessi in sequenze differenti e ciò permette di trarre un significativo guadagno dalla comparazione delle due informazioni, dopo una appropriata riorganizzazione dei dati stimati.

Un ulteriore guadagno è ottenuto reiterando le stime più volte, usando alternativamente i valori stimati dai due decoder, fino a quando viene deciso in modo definitivo (*hard*) se al bit ricevuto è assegnato il valore 0 oppure 1.

I turbo codici sono utilizzati, oltre che nel sistema UMTS, dallo standard DVB-RCS.

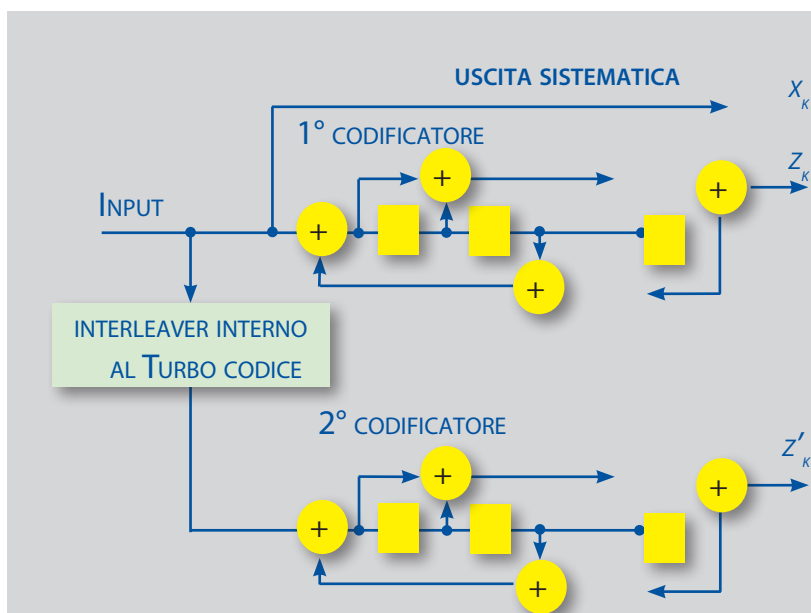



Fig. 4 - Codificatore utilizzato per i sistemi UMTS: segue i criteri di progetto indicati nel documento originale di Berrou e Glavieux del 1993 e utilizza una coppia di semplici codificatori convoluzionali identici. Per ogni bit in ingresso x_k vengono generati un bit x_k (il codice è sistematico) e due bit di parità z_k e z'_k: il code rate *R* è 1/3.

4. I codici LDPC

Con la scoperta dei Turbo Codici si avviò una rivalutazione degli schemi di codifica proposti nel passato, caratterizzati da una ridotta complessità, sfruttatabile per realizzare schemi di decodifica iterativi. Nell'ambito di tale analisi degli schemi precedentemente trascurati fu ripreso il lavoro iniziato con la tesi di dottorato di Robert Gallager nel 1962, sui codici LDPC [18].

In quanto codici lineari a blocco, i codici LDPC possono essere rappresentati mediante matrici (si veda come esempio il codice di Hamming, primo capitolo, § 4): la matrice per il calcolo dei bit di parità H e la matrice generatrice G. Per migliorare l'efficienza del codice, la matrice H deve essere costruita in modo che la distanza minima sia la più grande possibile: ciò implica che la matrice sia "sparsa", ovvero gli 1 siano in numero ridotto rispetto agli 0, da qui la denominazione di codice LDPC, cioè a bassa densità dei bit di parità. La matrice H è caratterizzata da n colonne e da n-k righe, dove k sono i bit di parità.

Se il numero di 1 presenti sulle colonne è costante e se il numero di 1 presenti sulle righe è costante il codice LDPC si dice regolare, altrimenti è detto irregolare. In figura 5 è un esempio molto semplice di codice LDPC con una matrice H (20,15).

I codici LDPC hanno il vantaggio di offrire prestazioni prossime a quelle teoriche (limite di Shannon), di

Fig. 5 - L'esempio proposto da Gallager nel 1963. E' basato su una matrice H (*parity check matrix*) costituita da 20 colonne e 15 righe caratterizzata da avere 4 elementi a 1 per ciascuna riga e tre elementi a 1 per ciascuna colonna.

essere adatti a differenti tipi di canale e di richiedere tempi di decodifica che crescono linearmente con le dimensioni del blocco. Inoltre sono realizzabili schemi che permettono un elevato grado di parallelismo sia in fase di codifica che di decodifica.

Sono stati proposti diversi algoritmi per costruire matrici H, anche utilizzando schemi di generazione pseudo casuali. Non è un problema complesso quello di generare codici LDPC con buone prestazioni e quelli con prestazioni migliori sono di tipo irregolare. La difficoltà di progettazione consiste nel mantenere bassa la complessità del codificatore e del decodificatore.

Fig. 6 - In pratica la matrice H deve essere di grandi dimensioni ed è importante definire la sua struttura in modo da minimizzare la complessità del codificatore e del decodificatore, anche in termini di memoria per rappresentarla. Nell'esempio la disposizione degli 1, rappresentati dalle linee diagonali, è tale da consente un risparmio di memoria e la possibilità di individuare una struttura in sottomatrici. Questa matrice è costituita da 16200 colonne e 5400 righe ed è adottata nei sistemi DVB di seconda generazione.

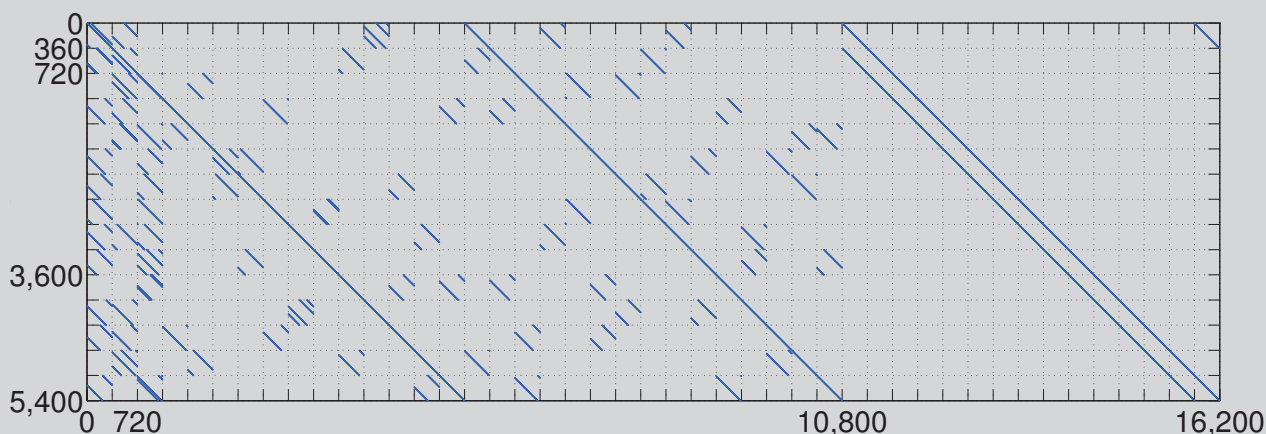
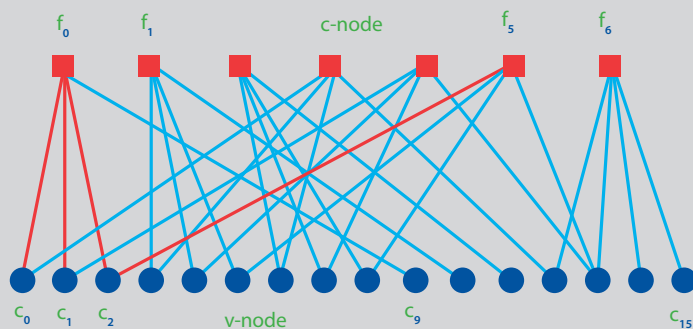



Fig. 7 - Nella rappresentazione proposta da Tanner il codice è descritto da un grafico bipartito.
Consiste di due tipi di nodi: i v-node e i c-node.
I v-node (*variable-node*) sono di solito indicati con la lettera *c* e sono in numero pari al numero di bit che costituiscono la parola di codice, ovvero *n*, quante le colonne della matrice H. I c-node (*check node*) sono indicati con la lettera *f* e sono tanti quante le righe di H, ovvero *n-k*. Il c-node f_i è connesso con il v-node c_j solo se l'elemento h_{ij} della matrice H è pari a 1.
L'algoritmo di decodifica prevede che ciascun nodo *f* riceva i contributi dai nodi *c* ad esso collegati e determini i valori di parità ritenuti corretti. Il calcolo può essere *hard*, semplice XOR dei contributi, oppure di tipo *soft*. Nel passo successivo tali valori sono inviati ai nodi *c*, e per ciascuno di essi è determinato, in base al valore precedente e ai contributi dei nodi *f* connessi, il nuovo valore di *c*.
Il processo è iterativo, ha termine quando non vi sono più variazioni nei valori di *c*, e quindi la parola di codice calcolata è ritenuta corretta, oppure quando si raggiunge il numero massimo previsto di iterazioni.
Una scelta opportuna della disposizione degli 1 nella matrice H, e quindi dei collegamenti fra i nodi del grafico, consente di effettuare i calcoli in parallelo, diminuendo il tempo di latenza dovuto alla decodifica.
Nell'esempio riportato, la matrice H in alto e il grafico di Tanner in basso corrispondono allo stesso codice, molto semplice. Il nodo f_0 è calcolato in base ai contributi c_0, c_1, c_2 e c_9 e, nel passo successivo, contribuisce, insieme a f_5, a determinare il nuovo valore di c_2...

Infatti le prestazioni dei codici migliorano al crescere delle dimensioni del blocco, come già era stato evidenziato da Elias nel 1955 [14]. I codici LDPC possono avere prestazioni migliori dei Turbo Codici quando la lunghezza del blocco è elevata, dell'ordine di alcune decine di migliaia di bit (figura 6)

Una rappresentazione alternativa del codice è quella grafica, introdotta da Tanner (figura 7). Se la matrice H è scelta senza alcuna restrizione, le connessioni, rappresentabili con il grafico di Tanner, appaiono a distribuite a caso e l'accesso alle n connessioni implica n cicli di clock. E' quindi vantaggioso adottare strutture di H tali da consentire una parallelizzazione parziale, in modo che un certo numero di nodi venga processato in parallelo.

Così come per i Turbo codici, anche per quelli LDPC il guadagno in termini di prestazioni è ottenuto grazie alla decodifica iterativa. Il numero di iterazioni è elevato (almeno 30) e il numero di calcoli, seppur semplici, per ciascuna iterazione è proporzionale alle dimensioni della matrice, pertanto la complessità totale della decodifica è superiore a quella richiesta per i Turbo Codici. Al crescere del numero di iterazioni, cresce la quantità di memoria necessaria e si incrementa la latenza (il ritardo nella decodifica).

Anche nel caso dei codici LDPC le realizzazioni pratiche utilizzano la decodifica *soft*, che consente una più rapida convergenza dell'algoritmo di decodifica.

I codici LDPC sono stati adottati per la prima volta in uno standard dal gruppo DVB-S2 [22]. Lo schema scelto per il sistema di diffusione via satellite di seconda generazione DVB-S2, è stato successivamente adottato per il DVB-T2 per il terrestre [27] e DVB-C2 per la distribuzione via cavo

La struttura prescelta è basata sulla concatenazione di due codici: un BCH come codice esterno e un LDPC come codice interno.

Il codice LDPC utilizzato dal DVB è denominato *Extendet IRA code*, ha il vantaggio di una ridotta (lineare) complessità del codificatore. La lunghezza di parola può essere n=64800 (per le trame normali) o n=16200 (per le trame corte). Per comprendere la complessità di decodifica, si consideri che, con la lunghezza di 64800 bit, ad ogni iterazione è necessario accedere e calcolare circa 300000 dati e il numero di iterazioni per garantire le prestazioni è pari a 30.

Per ridurre la complessità e la latenza nelle fasi di codifica e decodifica, i codici LDPC adottati dai sistemi DVB di seconda generazione sono caratterizzati da "matrici di parità" H sparse, disposte in modo da consentire l'individuazioni di sottomatrici, su cui operare in parallelo (figura 6).

L'uso dei codici LDPC permette un'elevata flessibilità nella scelta del *code rate,* infatti *R* può assumere per il DVB/S2 i valori 1/4, 1/3, 2/5, 1/2, 3/5, 2/3, 3/4, 4/5, 5/6, 8/9, e 9/10; mentre nel caso del DVB/T2 i valori possibili sono 1/2, 3/5, 2/3, 3/4, 4/5, 5/6.

Altri standard che utilizzano questi codici sono le nuove versioni di WiMax mobile e WiFi. Anche il sistema di televisione terrestre DTMB della Repubblica Popolare Cinese adotta uno schema basato su BCH e LDPC, con una lunghezza di 7488 bit.

5. Mascheramento

Nel primo capitolo si è visto che l'adozione dei codici serve a rivelare la presenza degli errori e, normalmente, a correggerli riducendo la probabilità di errore residua a valori tali da rendere non percepibili o accettabili gli effetti dovuti agli errori non corretti.

In § 5 del primo capitolo si è accennato agli effetti del superamento della capacità correttiva del codice: in ricezione si passa da una condizione ottimale, in cui tutti gli errori sono corretti, ad una condizione in cui, anziché recuperare integralmente l'informazione, vengono introdotti ulteriori errori e si perviene rapidamente ad una condizione di non funzionamento, di non-servizio.

Nel caso in cui il servizio lo consenta, si richiede la ritrasmissione dell'insieme di dati non ricevuti correttamente: ciò è possibile se è disponibile un canale di ritorno e se i dati possono essere ritrasmessi, ad esempio nel caso in cui si tratti di file di dati prodotti da un server. Oppure si attende una nuova ricezione degli stessi dati, nel caso di ritrasmisisone ciclica, come avviene per alcuni servizi quali il Televideo, esempio riportato nel primo capitolo.

Nel caso della diffusione di informazioni audio e video, in broadcasting o streaming, si può invece sfruttare la ridondanza residua, ancora presente nell'informazione ricevuta anche quando siano state utilizzate efficienti tecniche di compressione e codifica.

Le strategie di correzione e di interleaving consentono non solo di ridurre, grazie ad esempio alla concatenazione di codici, il numero di errori residui, cioè quelli che superano le capacità correttive del codice, ma anche di fornire una valutazione dell'attendibilità dei dati, in particolare se si adottano tecniche di decisione *soft*, al fine di segnalarla allo stadio di decodifica successivo.

Al momento in cui si estraggono dai dati associati al flusso binario le informazioni relative ai pixel d'immagine o ai campioni audio, è quindi nota l'affidabilità di tali informazioni e, se è elevata la probabilità che siano erronee, è possibile ricorrere alla tecnica del mascheramento (*concealment*).

Ad esempio, nel caso del segnale video, è possibile utilizzare la correlazione con i pixel contigui o fra blocchi e macroblocchi (nel caso di codifica MPEG) contigui nel tempo (appartenenti ai quadri precedenti o successivi) e ricostruire l'informazione mancante con pixel, blocchi o macroblocchi stimati a partire da quelli presenti in memoria e ritenuti corretti. Nel caso di perdita di gran parte dell'informazione (ad esempio in presenza di fading o rumore impulsivo) il mascheramento avviene congelando l'intera immagine presente in memoria, fino a quando la decodifica riprende correttamente.

Le tecniche di mascheramento di solito danno origine a difetti percepibili dall'utente, ma meno fastidiosi rispetto all'interruzione completa del servizio o alla visualizzazione di porzioni di immagine completamente scorrelate.

6. Cancellazioni (erasure)

Una sempre maggior mole di traffico dati si attua sulla rete Internet. Il protocollo internet (IP) prevede l'organizzazione dei dati in pacchetti dotati di una intestazione (*header*) che racchiude l'indirizzo la sorgente e la destinazione del pacchetto e spesso anche un numero che ne indica la posizione assoluta o relativa all'interno della sequenza che compone il flusso (*stream*) di dati. I pacchetti vengono instradati, seguendo i percorsi ritenuti più opportuni, attraverso la rete fino a raggiungere la destinazione. A volte, per le ragioni più varie (ad esempio *overflow* dei *buffer* presenti nei *router* intermedi), alcuni dei pacchetti non raggiungono la destinazione oppure i pacchetti ricevuti non vengono considerati validi, perché sono rivelati errori non correggibili.

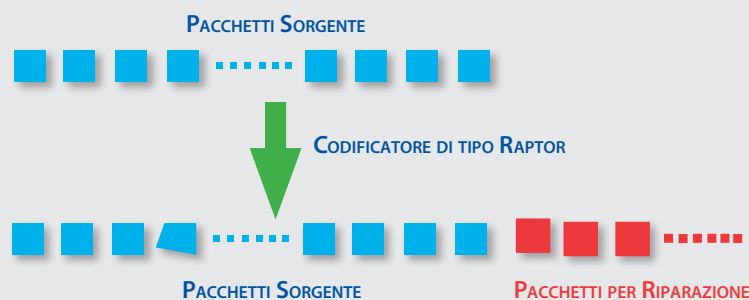
In questi casi si potrebbe procedere con la richiesta di ritrasmissione dei pacchetti mancanti, cioè cancellati (*erasure*), ma di fatto tale approccio può risultare non praticabile a causa della distanza fra il ricevitore e la sorgente (e quindi del tempo di latenza), della complessità o tipologia della rete (canali *wireless* o via satellite) o dal tipo di sorgente (un server che deve servire contemporaneamente più utenze e quindi non è in grado di gestire le richieste di ritrasmissione).

I *fountain code* sono codici che offrono una soluzione: la sorgente invia l'informazione in modo ridondante, tale da consentire la sua ricostruzione anche da parte del ricevitore soggetto ad un numero di erasure non superiore a quello massimo prevedibile (*worst case*). Il nome *fountain* è rappresentativo del concetto della fontana (la sorgente) che riempie il bicchiere (il ricevitore) anche nel caso in cui parte dell'acqua (i pacchetti dati) prodotti dalla fontana va persa.

Il *fountain code* produce, a partire da k simboli, un flusso teoricamente illimitato di simboli; il decodificatore è in grado di recuperare i k simboli a partire da un insieme di n simboli ricevuti; il codice ha buone prestazioni se n è prossimo a k, e se il tempo di decodifica è direttamente proporzionale a k. Il simbolo è, genericamente, un vettore di bit, ad esempio un pacchetto dati.

I codici LT sono stata la prima classe di *fountain code* utilizzabile in pratica. I *Raptor code* [28] sono una delle classi di *fountain code* caratterizzati da un tempo di codifica e decodifica lineare rispetto a k. Questi codici hanno due stadi di codifica: un *pre-code* o *outer-code* e un *inner-code*. Il *pre-code* può, a sua volta, essere la concatenazione di due codici, ad esempio un codice di Hamming e un LDPC. Il codice interno è un codice LT.

Alcuni dei più recenti standard quali il MBMS nell'ambito del 3GPP per la telefonia mobile di terza generazione e DVB-H [29] e DVB-SH per il datacast

Fig. 8 - I dati che costituiscono i file da trasmettere sono generalmente protetti da FEC, a livello fisico. Il file è organizzato in un numero fisso di *source symbol*; a livello applicazione il codificatore *Raptor* genera, oltre ai *source symbol,* un numero di *repair symbol* variabile in funzione del numero di *erasure* previsti nelle condizioni peggiori, cioè di pacchetti che si ritiene probabile vadano persi. Se è disponibile un canale di ritorno, si possono anche prevedere servizi in cui vengano generati, su richiesta, ulteriori *repair symbol*, destinati agli utenti che non sono stati in grado di ricostruire l'intera informazione. Lo scopo è quello di adattare la banda utilizzata, minimizzando quella necessaria a contrastare le *erasure* e riducendo al minimo la latenza, almeno per gli utenti che godono di migliori condizioni di ricezione.

verso i dispositivi mobili adottano il protocollo FLUTE che prevede, opzionalmente, l'uso dei codici *Raptor*.

I codici *Raptor* previsti in FLUTE *sono* di tipo sistematico, cioè i simboli del messaggio originale sono compresi fra quelli ricevuti. Si prevede la mappatura dei file in simboli denominati simboli sorgente (*source symbol*) e la generazione di simboli aggiuntivi utilizzati per la riparazione (*repair symbol*) una appropriata strategia basata sull'adozione di codici per la correzione degli errori consente di minimizzare sia l'effetto della perdita di pacchetti sia l'occupazione di banda (figura 8).

7. Oltre il limite?

Nel mese di luglio 2010 è stato diffuso un comunicato stampa su un nuovo dispositivo SoC basato su LDPC per gestire la memorizzazione dei dati su hard-disk. In particolare esso trova applicazione per i dischi da 2,5 pollici, il segmento a crescita più veloce nel mercato degli hard-disk, portando la capacità a 320 GB e quindi consentendo di continuare l'andamento nella crescita di capacità di memorizzazione, che raddoppia ogni 18 mesi.

Rappresenta l'ultimo progresso, in ordine di tempo, che possiamo attribuire al contributo essenziale delle tecniche di protezione degli errori.

Le tappe principali della storia che ha portato alle prestazioni finali dei sistemi attuali sono riassunte in figura 9.

Abbiamo visto nel secondo capitolo che è una storia che ha un inizio ben definito, temporalmente: coincide con la la pubblicazione dell'articolo di Shannon nel 1948. Fin dall'inizio la storia introduce due dei principali protagonisti e indica il possibile lieto fine. I protagonisti sono lo stesso C. E. Shannon, e R.W. Hamming, il cui codice è citato in tale articolo come esempio di codice "efficiente". E il lieto fine è costituito dal raggiungimento del limite di Shannon.

Negli anni '50, il computer, appena nato, è il primo campo di applicazione. Hamming inventa il codice proprio per rendere possibile il funzionamento del computer, che altrimenti, a causa degli errori, si blocca in continuazione. Il codice RS (Reed -Solomon)è del 1960; Reed raggiunge la notorietà per aver realizzato un computer compatto, delle dimensioni di una scrivania.

Negli anni '60 e '70, l'attenzione si focalizza sulle comunicazioni spaziali; il Jet Propulsion Laboratory e la NASA diventano il punto di aggregazione degli esperti che collaborano per mettere a punto i sistemi che consentano l'acquisizione delle informazioni raccolte dalle missioni spaziali. Protagonisti sono A.J. Viterbi e I.S. Reed. Il codice Reed-Muller, l'algoritmo di Viterbi per la decodifica dei codici convoluzionali e infine i codici concatenati RS e convoluzionali con decodifica di Viterbi sono le tappe che permettono l'esplorazione dei pianeti del sistema solare a partire dal 1969 fino alla fine del secondo millennio.

Tornando al campo della memorizzazione: gli schemi basati su codici prodotto RS trovano un ampia diffusione sia per la registrazione su supporto magnetico (disco o nastro) che su disco ottico. Sia il CD (1982) che il DVD (1996) lo adottano.

E lo schema che consente le esplorazioni spaziali è anche alla base dei sistemi di diffusione televisiva digitale. Lo schema RS-interleaving-codice convoluzionale-decodifica di Viterbi è adottato per la prima trasmissione digitale sperimentale via satellite delle immagini in alta definizione durante i campionati mondiali di Italia '90 e successivamente dal DVB-S (1994), il sistema di diffusione televisiva via satellite, normalizzato dal gruppo presieduto da Mario Cominetti, del Centro Ricerche Rai.

Il codice RS all'inizio degli anni '90 è il protagonista assoluto nei tre campi (memorizzazione, comunicazioni spaziali, diffusione televisiva digitale) e sembra che i 3 dB che separano dal lieto fine, il raggiungimento del limite di Shannon, siano un ostacolo insormontabile, quando i Turbo codici, "inventati" nel 1993, rendono il limite a portata di mano.

Sono trascorsi pochi anni, necessari perché la novità venisse recepita da alcuni standard di telecomunicazioni (DVB-RCS, HSPA per la telefonia e WiMax), quando il gruppo che si occupa del successore del sistema di diffusione televisiva via satellite, il DVB-S2, presieduto da Alberto Morello del Centro Ricerche Rai, mette in competizione codici Turbo e LDPC.

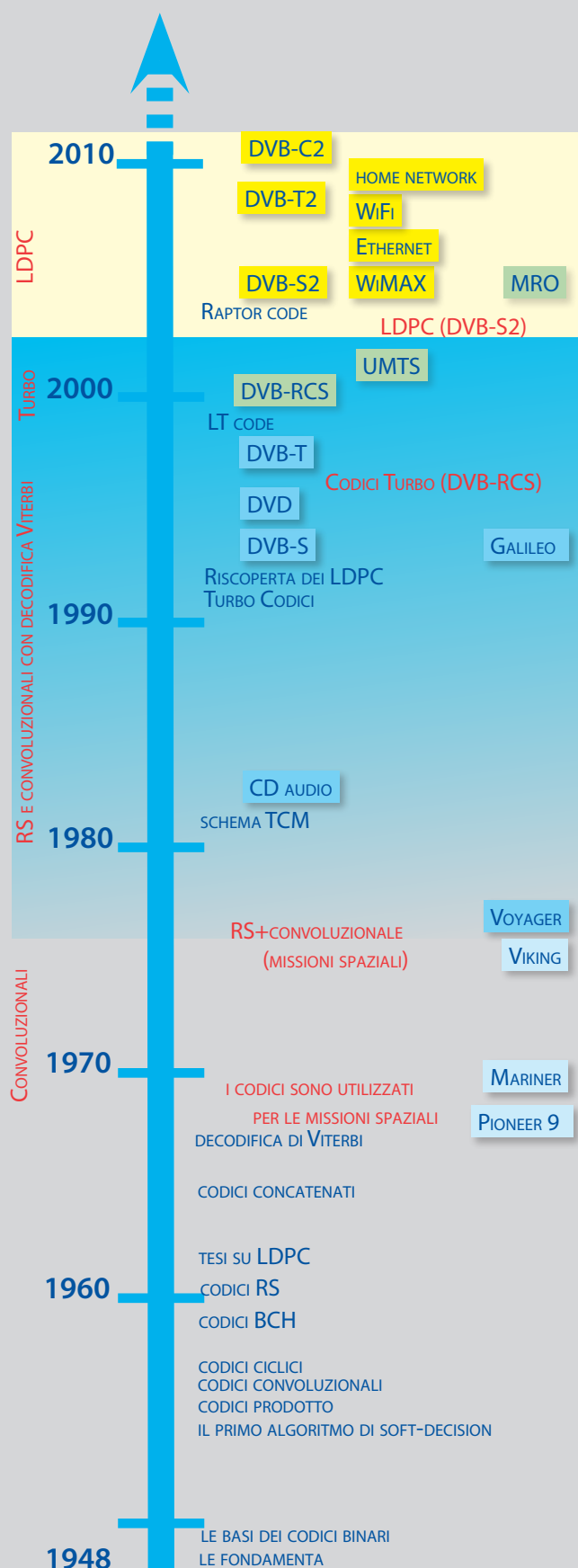



Fig. 9 - L'articolo di C. E. Shannon nel 1948 costituisce le fondamenta della teoria matematica dei codici e indica relazione fra capacità e banda di un canale soggetto a rumore gaussiano.

Gli anni '50 sono densi di importanti intuizioni: viene descritto il primo algoritmo di soft-decision (Wagner, 1954), sono inventati i codici prodotto e viene evidenziato che, per il raggiungimento del limite di Shannon, occorre utilizzare codici con elevata lunghezza di blocco (P. Elias, 1955), sono inventati i codici convoluzionali (P. Elias, 1955), i codici ciclici (E. Prange, 1957), e con la fine del decennio i codici BCH (A. Hocquenghem, 1959, e R.C. Bose e D.K. Ray-Chaudhuri, 1960).

Degli anni '60 sono le basi teoriche degli schemi di codifica oggi più diffusi: i codici RS (I. S. Reed e G. Solomon, 1960) e LDPC (R. Gallager, 1962), la concatenazione dei codici (G.D. Forney, 1965). Inizia l'uso dei codici per le missioni spaziali (Reed-Muller, 1969). I protagonisti in quegli anni collaborano al Jet Propulsion Laboratory della NASA, fra gli altri, Viterbi e Reed. L'algoritmo di Viterbi è del 1967.

Nel 1977, con le missioni Voyager, è introdotto lo schema basato sul codice RS concatenato con codice convoluzionale e decodifica di Viterbi. Inizia il lungo periodo di predominanza degli schemi basati su RS e codice convoluzionale.

Un ulteriore significativo passo di avvicinamento al limite è ottenuto applicando la codifica convoluzionale ai simboli della modulazione: è lo schema Trellis Coded Modulation (Ungerboeck, 1982).

Le tecniche di integrazione su larga scala rendono finalmente possibile l'adozione degli schemi di decodifica anche su blocchi di elevata lunghezza, guadagnando quindi in efficienza. Questi schemi possono essere adottati negli standard destinati a prodotti per il grande pubblico. Il primo è il CD audio (1982), che adotta uno schema basato sul prodotto di codici RS. Analogo schema, ma con migliori prestazioni, caratterizza il DVD (1996).

Agli inizi degli anni '90 il limite non è ancora stato raggiunto, ma i risultati degli ultimi 15 anni di sviluppo e realizzazioni sembrano stabili. Uno schema efficiente e collaudato è quello adottato per le missioni spaziali. E' un'ottima ragione per utilizzarlo anche per il nascente standard di diffusione da satellite, DVB-S (1994).

L'evento inatteso, che riduce la distanza che ancora separa dal limite è l'invenzione dei Turbo codici (C. Berrou, A. Glavieux, P. Thitimasjshima, 1993). I turbo codici vengono adottati dagli standard DVB-RCS e UMTS (ETSI TS 125 212, 2001). In campo spaziale è utilizzato dal MRO (Mars Reconnaisance Orbiter, 2005).

Nel 1994 sono riscoperti gli LDPC.

Sono "la soluzione finale", vengono adottati per le nuove generazioni di standard per le comunicazioni: DVB-S2 (2005), WiMAX (IEEE-80216e, 2005), 10GBase-T Ethernet (802.3an, 2006), WiFi (IEEE 820.11n, 2007), DVB-T2 (2008), G.hn (ITU G.9960, 2009), DVB-C2(2010).

Vincono gli LDPC, e in breve tempo il successo degli schemi basati su LDPC si estende alle nuove generazioni di standard: dopo il DVB-S2 (2005), è adottato per gli altri standard televisivi DB-T2 per la diffusione terrestre e DVB-C e per la distribuzione via cavo, e dagli standard per i collegamenti a microonde Wi-MAX, per le reti wireless WiFi, Ethernet 10GBase-T, per la rete domestica G.hn con distribuzione su linee elettriche, telefoniche e coassiali fino a 1 Gbit/s (ITU G.9960, 2009).

Il loro uso si afferma anche a livello di protezione dei servizi per downloading e streaming (protocollo FLUTE).

E infine, come abbiamo visto all'inizio, arrivano anche a sostituire i sistemi basati su RS anche per i sistemi di memorizzazione.

Il comunicato stampa citato proclama l'attualità della nuova architettura basata su LDPC, in sostituzione su quella basata su codici RS, inventata quasi 50 anni fa.   E' vero: l'articolo di Reed-Solomon è esattamente di 50 anni fa, del 1960, ma quello di Gallager, che descrive i codici LDPC, è del 1962, quarantotto anni fa.

Ai risultati di oggi hanno contribuito tutti i protagonisti della nostra storia. Il limite di Shannon è raggiunto da codici con elevata lunghezza blocco, come preconizzava Elias nel 1955, massimizzando la distanza di Hamming fra le parole di codice, utilizzando la soft-decision, adottando schemi che consentano di effettuare la decodifica in parallelo. Ma soprattutto il limite è raggiunto dagli schemi attuali perché oggi la densità di integrazione e la velocità di calcolo consentono di integrare memoria necessaria e algoritmo in una piccola superficie di silicio: in definitiva di realizzare un SoC.

La nostra storia è dunque una storia a lieto fine, il limite indicato nel 1948 è raggiunto, dopo circa mezzo secolo.   Eppure, come è tipico nell'ambito scientifico, l'obiettivo non è il raggiungimento del limite, ma è il suo superamento, lo scoprire ciò che ci attende, oltre l'orizzonte.

Quale sarà la continuazione di questa storia, il progresso dovuto ai codici per la protezione contro gli errori, nel campo della memorizzazione, delle esplorazioni spaziali e delle comunicazioni sulla terra?

Bibliografia

1. M. Barbero, N. Shpuza: "Le origini del video digitale (la rac. ITU-R BT.601)", Elettronica e Telecomunicazioni, aprile 2003
2. M. Barbero, N. Shpuza: "I format HDTV (le rac. ITU-R BT.709 e BT.1543)", Elettronica e Telecomunicazioni, aprile 2005
3. M. Barbero, N. Shpuza: "Uno standard pervasivo (MPEG-2 video)", Elettronica e Telecomunicazioni, aprile 2003
4. M. Barbero, N. Shpuza: "Advanced Video Coding (AVC - H.264)", Elettronica e Telecomunicazioni, aprile 2003
5. M. Cominetti, D. Tabone:"televideo: la telematica in ogni casa", Rai, Direzione Commerciale, 1987
6. R. W. Hamming: "Error detecting and error correting codes", Bell Syst. Tech. J. 29:147-60,1950
7. I.S. Reed, G. Solomon: "Polynomial codes over certain finite fields", J. Soc. Indust. Appl. Math, pp. 300-304, 1960
8. "Intelligent RAID 6 Theory. Overview and Implementation", Intel whitepaper, 2006
9. C. E. Shannon: "A Mathematical Theory of Communication", The Bell System Technical Journal, vol. 27, pp. 379-423, 623-656, July, October, 1948.
10. M. J. E. Golay, "Notes on digital coding," Proc. IRE, vol. 37, p. 657, June 1949.
11. D. E. Muller, "Application of Boolean algebra to switching circuit design and to error detection," IRE Trans. Electron. Comput., vol. EC-3, pp. 6–12, Sept. 1954.
12. I. S. Reed, "A class of multiple-error-correcting codes and the decoding scheme," IRE Trans. Inform. Theory, vol. IT-4, pp. 38–49, Sept. 1954.
13. P. Elias, "Error-free coding", IRE Trans. Inform. Theory, vol. IT-4, pp.29-37, Sept. 1954.
14. P. Elias, "Coding for noisy channels," IRE Conv. Record, vol. 4, pp. 37–47, 1955.
15. E. Prange, "Cyclic error-correcting codes in two symbols," Tech. Rep. TN-57-103, Air Force Cambridge Research Center, Cambridge, MA, Sept. 1957.
16. A. Hocquenghem, "Codes correcteurs d'erreurs," Chiffres, vol. 2, pp. 147–156, 1959.
17. R. C. Bose and D. K. Ray-Chaudhuri, "On a class of error correcting binary group codes," Information and Control, vol. 3, pp. 68–79, Mar 1960
18. R.Gallager, "Low-density parity-check codes," IRE Trans. Information Theory, pp. 21-28, Jan. 1962.
19. A. J. Viterbi: "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm," IEEE Transactions on Information Theory, Volume IT-13, pages 260-269, April, 1967.
20. C. Berrou, A. Glavieux, P. Thitimasjshima: "Near Shannon limit error correcting coding and decoding: Turbo-codes", Proc. IEEE Int. Conf. on Commun., Geneva, maggio 1993, pp. 1064-1070.
21. D.J.MacKay, R.M.Neal: "Near Shannon limit performance of low-density-parity-check codes.", Elect. lett., vol. 32, pp. 1645-1646, August 1996.

22. A. Morello, V. Mignone: "Il sistema DVB-S2 di seconda generazione per la trasmissione via satellite e Unicast", Elettronica e Telecomunicazioni, dicembre 2003.

23. M. Eroz, F.W. Sun, L.N. Lee: "DVB-S2 low density parity check codes with near Shannon limit performance", Int. J. Satell. Commun. Network., vol. 22, No 3, May-June 2004.

24. A. Morello: "Super Future: DVB-S2 Enables 140 Mps Super Hi-Vision By Satellite at IBC 2008", DVB Scene, No. 27, August 2008 (www.dvb.org)

25. N. Wells:"A Spec is Born. DVB-T2: A new Terrestrial Standard", DVB Scene, No. 27, August 2008 (www.dvb.org)

26. G.D. Forney, D.J. Costello: "Channel Coding: The Road to Channel Capacity", IEEE Procs, Vol. 95, Issue 6, p. 1150-1177, June 2007

27. *V. Mignone, A. Morello, G. Russo, P. Talone:* "DVB-T2 - la nuova piattaforma per la televisione digitale terrestre", Elettronica e Telecomunicazioni, Dicembre 2008.

28. *A. Shokrollahi*: "Raptor Codes",IEEE Trans. on Information Theory, vol. 52, No.6, June 2006

29. A. *Bertella, P. Casagranda, D. Milanesio e M. Tabone*: "Il sistema DVB-H per la TV Mobile", Elettronica e Telecomunicazioni, Dicembre 2005.

Le specifiche dei vari sistemi DVB possono essere reperite in:
www.dvb.org/technology/standards/index.xml

"Elettronica e Telecomunicazioni", nata nel 1952 come "Elettronica e Televisione Italiana", è una rivista quadrimestrale di Rai Eri realizzata dal Centro Ricerche e Innovazione Tecnologica della Rai, sul cui sito è disponibile gratuitamente dal 2001.

Il Centro Ricerche e Innovazione Tecnologica (CRIT) della Rai nasce a Torino nel 1930 come "Laboratorio Ricerche" e dal 1960 ha sede in Corso Giambone 68. Successivamente assume la denominazione "Centro Ricerche" e, dall'ottobre 1999, quella attuale.
L'attività del Centro è coordinata dalla Direzione Strategie Tecnologiche.
Alla nascita, tra i suoi obiettivi ha la progettazione e realizzazione di impianti ed apparati di nuova concezione, non reperibili sul mercato. I profondi cambiamenti nello scenario delle telecomunicazioni hanno stimolato la trasformazione del Centro.
Ha ricevuto riconoscimenti a livello internazionale per i contribuiti forniti alle attività di studio e normalizzazione dei sistemi per la codifica dei segnali audio e video in forma digitale, allo sviluppo delle tecniche di compressione dei segnali attualmente alla base dei sistemi di trasmissione e registrazione dei segnali video, alla definizione degli standard di diffusione e trasmissione DVB.
Il Centro contribuisce all'evoluzione delle tecnologie relative al sistema radiotelevisivo e multimediale e supporta il Gruppo nelle scelte di indirizzo tecnologico e nella fase di sperimentazione e introduzione in esercizio di nuovi prodotti e sistemi. E' attivo in numerosi progetti finanziati in ambito europeo e nazionale e collabora con Università e Industrie per l'attività di ricerca, per la definizione dei nuovi standard e lo sviluppo dei nuovi servizi.

Rai Radiotelevisione S.p.A.
Centro Ricerche e Innovazione Tecnologica
Corso E. Giambone, 68 - I 10135 Torino
www.crit.rai.it